Journal of the
OPTICAL SOCIETY *of* AMERICA

VOLUME 37, NUMBER 3 MARCH, 1947

An Investigation of the Properties of Evaporated Metal Bolometers*

B. H. BILLINGS, W. L. HYDE,** AND E. E. BARR
Research Laboratory, Polaroid Corporation, Cambridge, Massachusetts

A general theoretical analysis is made to determine the variation of bolometer sensitivity with thickness and heat dissipation. To a first approximation it is shown that the best response at a given modulation frequency is obtained by making the bolometer as thin as possible and then by adjusting the dissipation constant so that

$$\gamma = ac\omega$$

where γ is the heat dissipation per unit area per unit temperature difference, a the thickness, c the heat capacity per unit volume, and ω the angular frequency of modulation. Experiments are discussed which show that this analysis is only partially correct because of the variation with thickness of the temperature coefficient of resistance. Measurements are presented of the variation of resistivity and temperature coefficient of resistance with thickness for nickel and bismuth. An interference method was used to determine the thickness. The range studied was from 190A to 760A. An investigation is also made of the variation of γ with frequency for bolometers cooled by conduction. A method is presented for predicting the general frequency response of bolometers with conductive cooling. A brief description is given of a bolometer which was built during the investigation and a comparison is made of this bolometer with several others which have been reported in the literature.

UP to the present time, bolometers have been used in general as laboratory tools for measuring radiation intensities in the infra-red. They were not easily available commercially and were built by the experimenter who planned to use them. Laboratory bolometers did not have to be unusually rugged. They could be shielded from vibration and individual measurements could last for several seconds.

The recent extensive use of bolometers in the field has made it necessary for them to be designed to meet a much more rigid set of specifications. They had to withstand the jars and shocks of military equipment and had to operate successfully in an intense sound field. They had to detect radiation of the order of a few ergs per second and had to perform this detection in a few milliseconds. In addition to these requirements, the bolometers had to be capable of being produced in large quantities.

Although some of the analysis will be applicable to bolometers in general, this paper will be limited mainly to a discussion of the problem of obtaining adequate sensitivity and response time from bolometers made of evaporated metal films.

Although sensitivity and speed of response are mentioned as separate bolometer requirements, these two properties are very intimately connected. In any treatment of radiation detectors,

* Presented at the Winter Meeting of the Optical Society of America, March 7–9, 1946, Cleveland, Ohio.
** Now at Harvard University, Department of Physics.

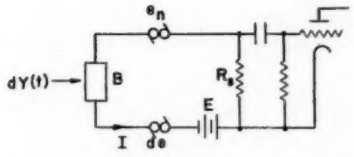

FIG. 1. The equivalent electrical circuit in which the bolometer is used.

it is necessary to consider this connection. The lowest amount of radiant power that can be detected by a given receiver depends on the noise generated in this receiver. The amount of this noise is generally proportional to the square root of the frequency band width of the system into which is sent the receiver signal. This band width is equal to the reciprocal of the longest time constant of the system, which may either be an amplifier, or a final galvanometer, or recorder, or the bolometer itself. The band width and the time taken by the instrument to make a reading are related by the equation

$$\Delta f = 1/2\pi\tau, \tag{1}$$

where Δf is the band width; τ is the time taken in the measurement. Thus the signal to noise is directly proportional to the square root of the measurement time. A receiver which in one piece of apparatus has a threshold of one erg per second may have a threshold in another piece of apparatus of 10^{-3} erg per second. The only difference between the two equipments may be that the second takes a million times as long to make a measurement. Laboratory bolometers were usually designed to work into galvanometers of the order of several seconds time constant. These same bolometers would have much smaller thresholds if used in a system which allowed only a few milliseconds per measurement.

The fact that early bolometers were used in narrow band systems does not necessarily mean they would not have been adequate in wide band systems. For pulse detection, a bolometer of a given time constant will give maximum sensitivity when operated with one definite amplifier passband. With an amplifier band starting at zero frequency, and a pulse of energy whose length is equal to the bolometer time constant, defined as $1/2\pi f$, where f is the bolometer high frequency cut-off, the best response is obtained with a bolometer whose high frequency cut-off coincides with the amplifier cut-off. This follows from a calculation of the Fourier spectrum of the pulse. If the amplifier cut-off is at a lower frequency than the bolometer cut-off, some of the high frequency components of the pulse are lost and the signal/noise ratio is decreased. On the other hand, if the amplifier cut-off is higher than that of the bolometer, more noise is passed than necessary and again the signal/noise ratio is less.

With this background we will be able to assess properly the characteristics of any bolometer with respect to the basic requirements.

APPROXIMATE SENSITIVITY EQUATION

We will assume the bolometer is being used in the electrical circuit shown in Fig. 1. The element B is in series with a battery of voltage E. The energy $dY(t)$ falling on the bolometer will generate a voltage de. The bolometer develops a statistical noise voltage e_n, where e_n is approximately equal to $Q(R\Delta f)^{\frac{1}{2}}$, where Q is a function of the absolute temperature of the bolometer, and Δf is the band width of the entire system which receives the bolometer signal, provided the system is linear throughout. For a given electronic system, e_n is essentially constant. R is the bolometer resistance, and R_s is the load resistance. We will assume, in addition, that the bolometer is so thin that heat conduction through its ends can be neglected compared to the heat loss from the surface by conduction and radiation.

For the heat balance equation of the bolometer element with no signal, and in a range of temperature near some reference temperature we can write

$$\beta\theta - I^2R - Y = 0, \tag{2}$$

where Y is the radiation from the bolometer surroundings, θ is the bolometer temperature, and β is the cooling constant of the bolometer. I is the bolometer current, and R is the bolometer resistance.

We now differentiate Eq. (2)

$$\beta d\theta - I^2dR - 2IRdI - dY = 0; \tag{3}$$

dY is the change in radiation on the bolometer and is called the signal. The differentials $d\theta$, dR, and dI are the change in temperature, resistance,

and current caused by this signal. For dR and dI we have

$$dR = R\alpha d\theta, \tag{4}$$

$$dI = -\frac{E}{(R+R_s)^2}dR, \tag{5}$$

where α is the temperature coefficient of resistance of the bolometer.

We also have

$$de = EdR/(R+R_s), \tag{6}$$

where de is the bolometer equivalent generator voltage which corresponds to the heat signal dY. If we combine Eqs. (3)–(6) we get

$$\frac{de}{dY} = \frac{R}{R+R_s}\cdot E\alpha \Big/ \left[\beta - \frac{EI\alpha R}{R+R_s}\left(1-\frac{2R}{R+R_s}\right)\right]. \tag{7}$$

For the metal bolometers discussed in the paper, β was of the order of magnitude of 10^{-3} whereas $EI\alpha$ was of the order of magnitude of 10^{-5}. It was thus possible to neglect the last term in the denominator of Eq. (7) without having more than a few percent error. We replace $R/(R+R_s)$ by the letter C and rewrite Eq. (7) as

$$de/dY = CE\alpha/\beta. \tag{8}$$

To obtain the response when dY varies with time we must consider h, the bolometer heat capacity. Instead of making dY a general function of the time, we can operate with the function $dY\cos\omega t$ and in the case of a linear system build any other function either with a Fourier integral or Fourier series. Equation (3) thus becomes

$$hd(d\theta/dt)+\beta d\theta - I^2dR - 2IRdI - dY\cos\omega t = 0. \tag{9}$$

This equation can be carried through the same process as Eq. (3) to give

$$hd(de/dt)+\beta de + CE\alpha dY\cos\omega t = 0. \tag{10}$$

The steady-state solution of this equation is

$$\frac{de}{dY} = C\frac{E\alpha}{(h^2\omega^2+\beta^2)^{\frac{1}{2}}}. \tag{11}$$

This expression will be called the AC response.

In general, we are not interested in the response as defined above, but in the response divided by the statistical noise voltage e_n which we will call the sensitivity. We wish this sensitivity to be a maximum. From the equations we have developed, we can see the effect on the sensitivity of modifications of the various bolometer parameters such as length, width, heat capacity, dissipation, etc.

For the 200 to 600A thickness metal strips used in our program, we can assume that the dissipation constant β is a function only of the bolometer area. We can then write

$$\beta = \gamma lw, \tag{12}$$

where γ is the dissipation per unit area, l the bolometer length, and w the bolometer width. For the bolometer heat capacity h we can write

$$h = alwc, \tag{13}$$

where a is the bolometer thickness and c is the heat capacity per unit volume. R, the bolometer resistance, can be written

$$R = \rho l/aw, \tag{14}$$

where ρ is the bolometer resistivity at its operating temperature. The AC and DC sensitivities depend directly on the voltage applied to the bolometer. This voltage cannot be increased indefinitely. For the maximum voltage we can write

$$E = (\beta\varphi R)^{\frac{1}{2}}, \tag{15}$$

where φ is the maximum temperature to which the bolometer can be raised. Every bolometer has such a maximum temperature. The limitation may be bolometer burn-out, or change of α or β. In any case, there is a temperature above which the bolometer response does not improve. If we insert these values in Eqs. (8) and (11) the expressions for the AC and DC sensitivities can be calculated.

$$S_{DC} = \frac{\alpha}{Q\Delta f}\left(\frac{\varphi}{\gamma lw}\right)^{\frac{1}{2}}; \tag{16}$$

$$S_{AC} = \frac{\alpha}{Q\Delta f}\left(\frac{\varphi}{\gamma lw}\right)^{\frac{1}{2}}\left(\frac{\gamma^2}{a^2c^2\omega^2+\gamma^2}\right)^{\frac{1}{2}}. \tag{17}$$

The area of a bolometer is generally determined by the system in which it is to be used. In cases where there is freedom of choice in area, we see from these equations that for a given energy falling on the bolometer, the smaller bolometer will give the greater sensitivity.

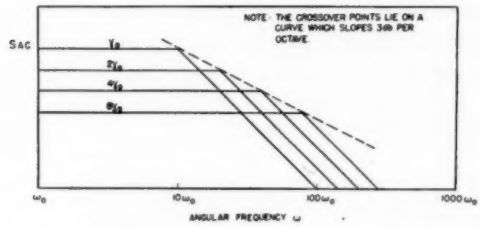

FIG. 2. The variation of the logarithm of the AC sensitivity of a bolometer strip with a change in the dissipation constant γ. The thickness and material are considered constant.

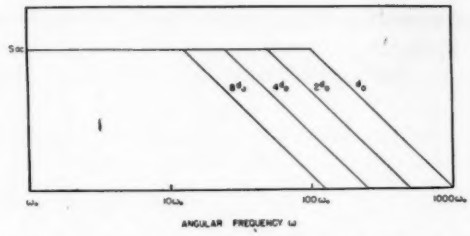

FIG. 3. The variation of the logarithm of the AC sensitivity of a bolometer strip with thickness. The dissipation constant and the material are kept constant.

The effect of a change in any one characteristic on the bolometer sensitivity can be seen from these approximate relationships. We see at once that the higher the temperature coefficient of resistance, the higher the bolometer signal/noise. Also the resistance of the bolometer does not enter the equations. Bolometer performance is not affected by resistance except perhaps through its correlation with α or φ. The behavior with respect to γ and a is not quite so simply expressed.

In Fig. 2 is shown a series of curves of bolometer response with all the parameters fixed except γ; as γ is increased, the bolometer response at high frequency is raised, but at the expense of the response at the low frequency end.

In Fig. 3 is shown the variation with a. Here the response at high frequencies is improved as the bolometer is made thinner. The low frequency response, however, is unaffected.

It is apparent that the highest response at a given frequency would be obtained by making the bolometer as thin as possible and by adjusting the dissipation constant to make

$$\gamma = ac\omega. \tag{18}$$

In order to make it possible to vary thickness from zero, it was decided to evaporate the strips on a thin backing of plastic. This technique had been used by Harris[1] for thermocouples and by Dewhurst[2] for sputtered bolometers. It was hoped that the plastic could be made sufficiently thin so that its addition to the bolometer heat capacity would be more than offset by the small thickness which could be obtained. The dissipation constant γ was not manipulated. The bolometers were operated in air at atmospheric pressure. In the final model, the plastic pellicle was supported on a glass base through which were sealed the metallic leads. The leads were connected to the bolometer by heavy evaporated copper contacts. Thus, it was not necessary to make any soldered connection directly to the bolometer strip on its plastic pellicle.

The details of the construction of the final model bolometer will be described elsewhere.[3]

VARIATION OF α AND ρ WITH BOLOMETER THICKNESS

During the period of design and testing of the final bolometer, experiments were done to select the best material and get the best value of thickness. Although we needed a bolometer with a time constant of 3 to 4 milliseconds, the thickness value could not be determined directly from Eqs. (16) and (17) mainly because so many of the constants in these equations were actually functions of the strip thickness. Accordingly, several materials were studied to determine the change with thickness of α and ρ, the temperature coefficient and the resistivity. One of the first problems in such an analysis is the determination of the meaning of the word "thickness." From electron microscope studies, such as those made by Picard and Duffendack,[4] it is apparent that evaporated films consist of a porous structure of individual metallic crystals or metallic agglomerates. If the surface of such a film has a roughness which is of the order of magnitude of the size of the individual crystals and if the number

[1] Louis Harris, Phys. Rev. **45**, 635 (1934).
[2] H. Dewhurst, Proc. Phys. Soc. **39**, 39 (1926).
[3] B. H. Billings, E. E. Barr, and W. L. Hyde, to be published.
[4] R. G. Picard and O. S. Duffendack, J. App. Phys. **14**, 291 (1943).

of crystal layers is small, it is difficult to determine just what is meant by the film thickness.

Any parameter used in measuring film thickness is dependent on thickness in a film of this type. If the film is weighed and bulk densities are assumed, the porosity of the film is ignored, and a thickness value is obtained which is quite different from the value which might be obtained by measuring the absorption of light through the film and assuming the bulk value of the absorption constant. Similarly, if the thickness of the film is determined by measuring its resistance and using bulk values of resistivity, a figure is obtained which also does not make any allowances for porosity.

The first technique was suggested by John Strong,[5] and is a modification of a method used by Tolansky[6] to sharpen the Newton interference fringes by half silvering the two surfaces between which the interference takes place. The general procedure was as follows:

A layer of metal of the same type used to make the bolometer was evaporated on a small optical flat. A one-half inch wide shield was placed across the coated flat and another layer of metal was then deposited with a thickness equal to that which we desired to measure. A second flat was then coated with a transparent coat of metal. When the metal coated surfaces of the two flats were placed in contact, a series of fine hairline fringes were observed. The fringes were observed both by reflecting monochromatic light from the surfaces and by sending monochromatic light through the sandwich. The light was obtained from an H-2 Mercury Lamp and a Wratten 77 filter. The optical flats were manipulated until the fringes were parallel straight lines perhaps $\frac{1}{4}''$ apart and normal to the boundary line between the thick and thin metallic deposits. In crossing this boundary line, the fringes were shifted a distance which depended on the increase in thickness of the air layer. This increase of thickness was a measure of what we call the interference thickness of the metallic film.

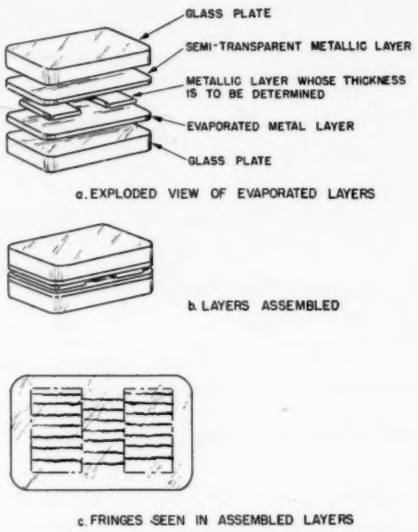
GLASS PLATE
SEMI-TRANSPARENT METALLIC LAYER
METALLIC LAYER WHOSE THICKNESS IS TO BE DETERMINED
EVAPORATED METAL LAYER
GLASS PLATE
a. EXPLODED VIEW OF EVAPORATED LAYERS
b. LAYERS ASSEMBLED
c. FRINGES SEEN IN ASSEMBLED LAYERS

FIG. 4. The method of measuring the thickness of a metallic film by interference.

In Fig. 4 is shown a sketch of the assembly. By transmission, the lines were, of course, sharper and brighter than by reflection. Without the Wratten filter, the silvered surfaces acted as an interference optical spectroscope and a separate fringe was formed by each line of the mercury spectrum. It was thus possible to work without the use of the Wratten filter and measure the

TABLE I. Values of the resistance thickness and interference thickness of films of nickel and bismuth.

Resistance of 6×1 mm strip (ohms)	Bismuth thickness measurements: Thickness a_I (angstroms)	Thickness a_ρ (angstroms)	Apparent ρ from a_I and resistance
204.5	760	352	2.50×10^{-4}
402.6	660	179	4.43×10^{-4}
604.8	570	119	5.74×10^{-4}
804.6	510	89	6.84×10^{-4}
Resistance of 6×1 mm strip (ohms)	**Nickel thickness measurements: Thickness a_I (angstroms)**	**Thickness a_ρ (angstroms)**	**Apparent ρ from a_I and resistance**
37	540	127	3.33×10^{-5}
71	340	66	4.01×10^{-5}
105	210	44	3.64×10^{-5}
141	190	33	4.46×10^{-5}

[5] John Strong, personal communication.
[6] S. Tolansky, Nature **153**, 314 (1944).

TABLE II. Measurements made by Columbani on thin films of nickel which were sputtered on glass.

a (angstroms)	ρ	α
15,000	6.29×10^{-5}	0.00439
3,800	$.266\times10^{-3}$	.0022
1,780	$.149\times10^{-2}$	.00005
380	2.8	−.0024

TABLE III. Variation of α with resistance of a 6×1 mm strip of evaporated metal.

Metal	Resistance in ohms of 6×1 mm strip $T=30°C$	$\alpha=\frac{\Delta R}{R\Delta T}$ $T=50°C$
Gold	bulk value	3.0×10^{-3}
	27.5	1.4×10^{-3}
	374	$.71\times10^{-3}$
Nickel	bulk value	4.3×10^{-3}
	22.5	2.6×10^{-3}
	91.0	1.2×10^{-3}
	296	1.1×10^{-3}
	1243	$.63\times10^{-3}$
Bismuth	bulk value	$+4.5\times10^{-3}$
	167	-1.7×10^{-3}
	710	-2.1×10^{-3}

thickness of the film with light of known wavelength.

The thickness was also determined by measuring the resistance of the film and using the bulk value ρ of the resistivity. In Table I are shown some values of the resistance thickness a_ρ and interference thickness a_I of films of nickel and bismuth.

In Table II are shown some measurements made by Columbani[7] on films of nickel which were sputtered on glass. He used the density thickness obtained by weighing his strips and using the bulk density. The tremendous difference between the two measurements of a at 380A in Tables I and II cannot be explained on the basis of the difference in thickness determination. Since his sputtered films were dark in appearance they were probably more porous than those which we obtained by evaporation.

The variation in α is shown in Table III.

It is clear that the value of α drops quite rapidly as the thickness is decreased. This is unfortunate as it indicates that although a bolometer can be made increasingly fast as its thickness is decreased, its sensitivity will drop because of the decrease in α. The fact that α finally becomes negative suggests immediately that, other things being equal, the best bolometer might perhaps consist of a material whose bulk value of α is negative. This is borne out by a brief study which we made of evaporated tellurium which has a bulk α of approximately -0.006. In thin films, this α was increased to -0.016. The strip with this value had a resistance of 10,000 ohms per square and since we were working with low impedance bolometers, this investigation was not continued.

It is interesting that Brockman[8] did not find a change in α or a change of resistivity with thickness. He prepared his nickel strips by electro-deposition and found identical values for resistance thickness and density thickness. His resistance thicknesses were 10–30 times the resistance thicknesses of the evaporated films prepared in this research. He may not have found a variation in α either because his films were so much thicker or because they were prepared by a completely different method.

It should also be pointed out that the structure of an evaporated metal film is very much a function of the type of surface on which it is evaporated. This is well illustrated by the work of Lovell[9] in which he produces large variations in film resistivity by evaporating onto contaminated surfaces.

VARIATIONS OF DISSIPATION CONSTANT WITH FREQUENCY OF MODULATION

Besides the variation of α and ρ with thickness, we investigated also the variation of γ, the dissipation constant, with modulation frequency.

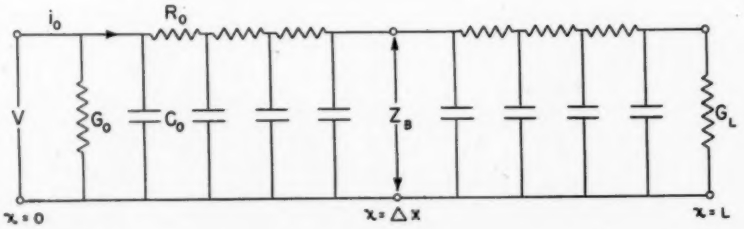

FIG. 5. The equivalent circuit of the bolometer and its support.

[7] Antoine Columbani, Ann. de Physique **19**, 272 (1944).
[8] F. G. Brockman, J. Opt. Soc. Am. **36**, 32 (1946).
[9] A. C. B. Lovell, Proc. Roy. Soc. **157**, 311 (1936).

This variation manifests itself whenever the cooling is conductive and results from the difference in velocity of heat waves of different frequencies. The analysis is carried through most easily by using a system of electrical analogies.

The conductive case has been treated in this manner for air by C. B. Aiken[10] and for solids by J. N. Shive.[11]

The bolometer and the various backing layers are represented in Fig. 5 as a series of long electrical lines. The equation for the voltage at the input can be written

$$V=\frac{i_0}{G_0+\left(\frac{jC_0\omega}{R_0}\right)^{\frac{1}{2}}\left(\frac{(R_0/jC_0\omega)^{\frac{1}{2}}\cosh(jC_0\omega R_0)^{\frac{1}{2}}\Delta x+Z_B\sinh(jR_0C_0\omega)^{\frac{1}{2}}\Delta x}{(R_0/jC_0\omega)^{\frac{1}{2}}\sinh(jC_0\omega R_0)^{\frac{1}{2}}\Delta x+Z_B\cosh(jR_0C_0\omega)^{\frac{1}{2}}\Delta x}\right)},$$

where V is the input voltage, i_0 the input current, R_0 the resistance of the line in ohms per centimeter, C_0 is the capacity in farads per centimeter; G_0 and G_L are the input and terminal conductances, and Z_B is the impedance of the remaining line in the system.

In place of the electrical circuit elements we can substitute the following thermal analogues, referring in each case to a unit area of bolometer:

The energy on the bolometer $dY(1+e^{j\omega t})$ is equivalent to the current $i_0(1+e^{j\omega t})$.

The surface temperature rise T_0 caused by the signal is equivalent to V_0, the voltage at the line terminals.

The thermal resistance per cm length—$1/k$ is equivalent to R_0, the electrical resistance per cm length.

The thermal capacity per cm length—C is equivalent to C_0, the electrical capacity per cm length.

The front face dissipation σ is equivalent to G_0, the input conductivity.

The rear dissipation σ' is equivalent to G_L, the terminal conductivity.

The thickness of the bolometer is Δx.

The thickness of the backing is L.

We also introduce the thermal diffusivity of the material, $K=k/c$.

The relative bolometer surface temperature then becomes:

$$T_0=\frac{dY}{\sigma+k\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\frac{\frac{1}{k}\left(\frac{K}{j\omega}\right)^{\frac{1}{2}}\cosh\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\Delta x+Z_B\sinh\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\Delta x}{\frac{1}{k}\left(\frac{K}{j\omega}\right)^{\frac{1}{2}}\sinh\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\Delta x+Z_B\cosh\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\Delta x}}, \tag{19}$$

where

$$Z_B=\frac{1}{k_B}\left(\frac{K_B}{j\omega}\right)\frac{\frac{1}{k_B}\left(\frac{K_B}{j\omega}\right)^{\frac{1}{2}}\sinh\left(\frac{j\omega}{K_B}\right)^{\frac{1}{2}}L+\frac{1}{\sigma'}\cosh\left(\frac{j\omega}{K_B}\right)^{\frac{1}{2}}L}{\frac{1}{k_B}\left(\frac{K_B}{j\omega}\right)^{\frac{1}{2}}\cosh\left(\frac{j\omega}{K_B}\right)^{\frac{1}{2}}L+\frac{1}{\sigma'}\sinh\left(\frac{j\omega}{K_B}\right)^{\frac{1}{2}}}.$$

Since these equations are quite complicated, they are reduced by Aiken and Shive to a series of equations valid at the extremes of the frequency range. This technique becomes difficult if the heat flow from the bolometer goes through several layers. A method for simplifying the analysis is described below.

We are not as interested in the exact shape of the bolometer response curve as in the frequencies at which the curve changes its general slope. These points can be located by making the approximation:

$$\tanh x=x/(1+x). \tag{20}$$

When the approximation is made, Eq. (19) can be expressed as the ratio of two polynomials. These polynomials can be factored, and produce

[10] C. B. Aiken, Electro-Mechanical Research Corporation, Houston, Texas (personal communication).

[11] J. N. Shive, Bell Laboratories (to be published).

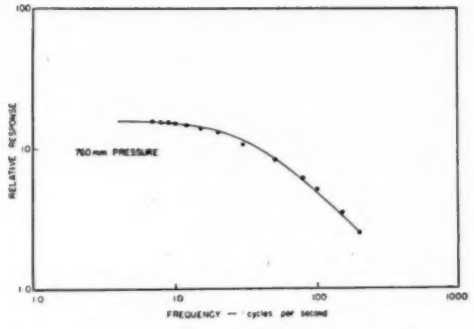

FIG. 6. The frequency response of a single strip nickel bolometer on a sandblasted glass disk.

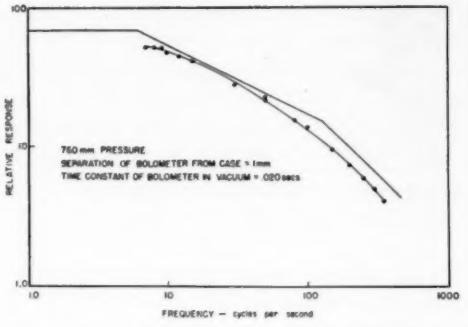
RELATIVE RESPONSE
760 mm PRESSURE
SEPARATION OF BOLOMETER FROM CASE = 1mm
TIME CONSTANT OF BOLOMETER IN VACUUM = .020 secs
FREQUENCY — cycles per second

FIG. 7. A comparison of the theoretical and experimental frequency response curve of single strip nickel bolometer.

roots in terms of the physical constants of the system. The general form of the equation will be

$$\frac{T}{dY}=S=\frac{A[(j\omega)^{\frac{1}{2}}+\omega_1^{\frac{1}{2}}e^{i\theta_1}][(j\omega)^{\frac{1}{2}}+\omega_2^{\frac{1}{2}}e^{i\theta_2}]\cdots[(j\omega)^{\frac{1}{2}}+\omega_l^{\frac{1}{2}}e^{i\theta_l}]}{[(j\omega)^{\frac{1}{2}}+\omega_{l+1}^{\frac{1}{2}}e^{i\theta_{l+1}}][(j\omega)^{\frac{1}{2}}+\omega_{l+2}^{\frac{1}{2}}e^{i\theta_{l+2}}]\cdots[(j\omega)^{\frac{1}{2}}+\omega_{l+n}^{\frac{1}{2}}e^{i\theta_{l+n}}]}. \tag{21}$$

S is now the relative response per unit voltage in the bolometer. The $(\omega_i)^{\frac{1}{2}}e^{i\theta_i}$ are the roots of the polynomials and have the dimensions of angular frequency to the one half power. The θ_i determine the relative phase of the temperature rise caused by different components of the input signal. A is a constant.

This expression is introduced for two reasons. First as a convenient way of determining the general shape of the bolometer response curve; second, as a means of determining the effect of a change in construction parameters on the shape of the response curve. To illustrate the use of Eq. (21) consider first the case of a bolometer which is radiation cooled. Equation (19) becomes

$$S=\frac{1}{\sigma+k\left(\frac{j\omega}{K}\right)^{\frac{1}{2}}\dfrac{\dfrac{1}{\sigma'}\dfrac{(j\omega/K)^{\frac{1}{2}}\Delta x}{1+(j\omega/K)^{\frac{1}{2}}\Delta x}+\dfrac{1}{k}\left(\dfrac{K}{j\omega}\right)^{\frac{1}{2}}}{\dfrac{1}{k}\left(\dfrac{K}{j\omega}\right)^{\frac{1}{2}}\dfrac{(j\omega/K)^{\frac{1}{2}}\Delta x}{1+(j\omega/K)^{\frac{1}{2}}\Delta x}+\dfrac{1}{\sigma'}}}. \tag{22}$$

Assume $(j\omega/K)^{\frac{1}{2}}\Delta x \ll 1$ and $(\sigma'\Delta x/k)\ll 1$. Eq. (22) becomes

$$S=\frac{1}{\sigma+\sigma'+(k\Delta x/K)j\omega}=\frac{1}{\sigma+\sigma'+jg\omega}, \tag{23}$$

where g is the heat capacity per unit area. This equation is now in the form

$$S=\frac{1/g}{\omega_0+j\omega}, \tag{24}$$

where $\omega_0=(\sigma+\sigma')/g$. Instead of plotting point by point the curve S *vs.* ω we draw on log-log paper the two asymptotes at $\omega\gg\omega_0$ and at $\omega\ll\omega_0$. The first is the line $S=1/g\omega$. The second is the line $S=1/g\omega_0$. The two lines intersect at $\omega=\omega_0$. These two straight line segments are sufficient to describe the behavior of the bolometer. The actual response at $\omega=\omega_0$ is 0.707 times the response at $\omega\ll\omega_0$, and the time τ, represented by the reciprocal of ω_0, is sometimes referred to as the bolometer time constant. In terms of physical parameters, we see that ω_0 can be shifted to higher frequencies either by reducing g or by increasing $\sigma+\sigma'$.

In Eq. (24) we considered ω_0 as the cross-over

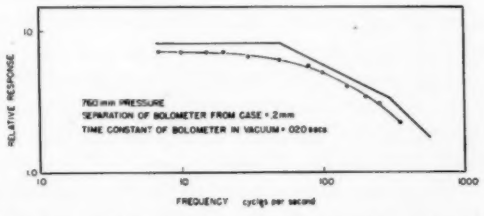
RELATIVE RESPONSE
760 mm PRESSURE
SEPARATION OF BOLOMETER FROM CASE = 2mm
TIME CONSTANT OF BOLOMETER IN VACUUM = .020 secs
FREQUENCY cycles per second

FIG. 8. A comparison of the theoretical and experimental frequency response curve of single strip nickel bolometer.

TABLE IV. The comparative characteristics of some infra-red receivers.

Element	Threshold by author	Conditions of measurements	Element time constant	Calculated threshold for 0.06 cm² area element with 30-cycle modulation frequency and 100-cycle band width electronic system. Element time constant unchanged	Calculated threshold for 0.06 cm² area element with 30-cycle modulation frequency and 100-cycle band width. Element developed by correct cooling to give maximum sensitivity
Thermistor bolometer[a]	2×10^{-8} watt	30-cycle band width system 0.003 sec. pulse on element of area 0.006 cm²	0.003 sec.	7.2×10^{-8} watt	7.2×10^{-8} watt
Dielectric bolometer[b]	3×10^{-8} watt	0.046-cycle band width. Steady illumination on element of area 1 cm²	3.5 sec.	2.3×10^{-4} watt	1.3×10^{-5} watt
Harris thermocouple[c]	2.2×10^{-10} watt	0.00053-cycle band width; 2 cycles sec. modulating frequency, 0.012 cm² area	.2 sec.	3.0×10^{-6} watt	6.9×10^{-7} watt
Polaroid bolometer	3.3×10^{-8} watt	100-cycle band width; 30-cycle modulated illumination on 0.06 cm² area element	.004 sec.	3.3×10^{-8} watt	3.3×10^{-8} watt

[a] J. A. Becker and W. H. Brattain, J. Opt. Soc. Am. **36**, 354 (1946).
[b] Steinhardt and Moon, J. Opt. Soc. Am. **28**, 148–162 (1938).
[c] Louis Harris, J. Opt. Soc. Am. **30**, 519 (1940).

of the two asymptotes to the response curve. We can consider then the ω_i's in the general equation as similar cross-overs. In this thermal system, the lowest frequency will be ω_{l+1}. In plotting the curve a series of straight lines are merely run between the points representing the response at the various cross-overs in ascending orders. The slope of the straight lines is determined by the power of a given ω_i and the slope of the proceding straight line.

The temperature $d\theta$ in Eq. (5) is the mean temperature. Equation (19) gives the surface temperature. At frequencies such that the heat wave-length in the strip is larger than the strip thickness, the mean temperature can be considered equal to the surface temperature. At higher frequencies, the mean temperature will be lower than the surface temperature. At very high frequencies the mean temperature of the strip will be proportional to $1/\sqrt{\omega}$ times the surface temperature.

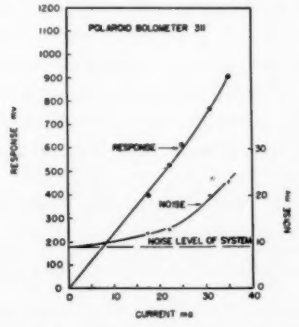

FIG. 9. The variation of bolometer response and noise with current.

For the temperature wave-length in the strip we can write

$$\lambda = 2\pi(2K/\omega)^{\frac{1}{2}}. \tag{25}$$

The cross-over of the two asymptotes will take place at the angular frequency at which

$$\Delta x \simeq \lambda/2\pi\sqrt{2}. \tag{26}$$

This is at a frequency

$$\omega = K/\Delta x^2. \tag{27}$$

Above this frequency we must add 3 db per octave to the slope of the bolometer response curve.

For bolometers whose thickness is of the order of ten microns, the structure is similar to the structure of the bulk material, and it is possible to use bulk values of the heat capacity in evaluating Eq. (21). For the metallic bolometers in this paper, the structure is different from that of the bulk material. The heat capacity is a function of thickness and must be especially evaluated. This evaluation was accomplished by measuring the bolometer's frequency characteristic in vacuum. We could rewrite Eq. (21) as

$$S = \frac{1}{\sigma + jg + 1/Z_B}, \tag{28}$$

where g is the heat capacity of the bolometer strip and its plastic backing, σ is the radiative dissipation, Z_B the impedance of the air space and case. In vacuum Z_B becomes infinite and Eq. (28) reduces to Eq. (24). By measurement of the cross-over frequency in vacuum and with a knowledge of the radiation constant, we can determine g.

The air space and the case have no unusual properties and we can use bulk values for their heat capacities and conductivities. When this air space is very small, the characteristics of the case become the contributing factor in the dissipation. We can consider the air layer as being in series with the bolometer and the case. When the layer is very thin, it can also be considered as a lumped impedance rather than a long line.

In Fig. 6 is shown the frequency response of a bolometer which behaves as though its dissipation were largely resistive in character. In Figs. 7 and 8 are a pair of bolometers whose frequency response departs quite widely from this characteristic. The only difference in these bolometers was in the thickness of the air layer separating them from the metal bolometer housing. The straight lines in each case represent the theoretical response obtained from Eq. (23).

Figures 6 and 8 illustrate not only the departure of the frequency curve from the simple cross-over type, but also illustrates the method of changing a bolometer's frequency characteristics by changing γ. The bolometer in Fig. 8 might be considered to have a shorter time constant than the bolometer in Fig. 7. This is the technique used by Langton.[12]

DEPARTURE FROM STATISTICAL NOISE

In addition to the approximation for variations in α, and ρ another approximation in Eq. (17) is that e_n is statistical and is proportional to the Johnson noise. For those bolometers whose interference thickness was much less than 200A, the measured noise was greater than theoretical. This extra noise was present only when current was passed through the element.

The variation of this extra noise with current seemed to depend on the bolometer material and on its past history. In Eq. (11) it is shown that the bolometer response is directly proportional to the voltage. If the factor of proportionality between the noise and the voltage is the same, the bolometer performance will not be improved by raising the voltage and the maximum sensitivity as given in Eq. (16) will not be reached. For some bolometers, the noise increases faster with an increase of voltage than the signal. For others, it increases at a slower rate. Variations with current of the noise and signal for a Polaroid bolometer which exhibited current noise are shown in Fig. 9.

Since the current noise was not in general constant over a period of weeks, those bolometers so afflicted were not considered satisfactory.

BOLOMETER PERFORMANCE

The performance of a typical bolometer is compared in Table IV with some other heat sensitive elements. In Columns 5 and 6 thresholds are calculated for the conditions demanded in this research. These thresholds represent the power required to give an rms signal equal to the rms noise. In Column 5 the manufactured characteristics of the elements are used. In Column 6 it is assumed that the various elements are re-designed with conductive cooling chosen to give maximum response at 30 cycles.

It is unfortunate that many elements mentioned in the literature are not sufficiently well described to enable sensitivity calculations to be made.

The authors wish to express their appreciation to Mrs. J. R. Platt, Mrs. L. A. Weaver, Jr., and Miss E. E. Comerford for their continued help during the program, and to members of the Polaroid Corporation, and especially to Mr. E. H. Land, for many stimulating discussions.

[12] W. G. Langton, J. Opt. Soc. Am. **36**, 355 (1946).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

The Measurement of the Intensity Ratios of Spectral Lines with Electron Multiplier Photo-Tubes

K. G. KESSLER AND R. A. WOLFE
Randall Laboratories of Physics, University of Michigan, Ann Arbor, Michigan

(Received September 23, 1946)

Some properties of electron multiplier photo-tubes are investigated. The use of these tubes in quantitative spectrographic analysis is discussed, and a scheme for reducing the effect of the dark current is described. A sharply tuned electronic filter passing only the 120-cycle component of the signal is used to increase the signal to noise ratio. An amplifier circuit for measuring the intensity ratios of spectral lines is shown, and results from the application of this circuit to the analysis of chromium in steel are tabulated.

THE photographic plate as used for the quantitative measurement of spectral line intensities has placed a definite limit on both accuracy and speed in industrial spectrochemical applications. Methods of plate calibration are tedious and at best give only a fair approximation to the true intensity ratios. Other difficulties such as the non-uniformity of emulsions, the variation of plate sensitivity with humidity, and the errors introduced by the microphotometer restrict the accuracy of the photographic process. In industrial applications where speed must be considered, time consumed in processing, calibrating, and reading photographic plates is an important factor.

A direct reading method, by eliminating plate errors and time consumed by the photographic processes, solves many of the problems encountered in quantitative spectrographic analysis. Such a method must, however, have the desired sensitivity, and must not introduce errors greater than those of the photographic method.

Morris[1] has applied photo-sensitive Geiger-Müller counters to the problem, and the use of electron multiplier photo-tubes for direct reading work has been studied by Dieke,[2,3,4] Saunderson,[5] and others. This paper will deal only with the use of electron multiplier photo-tubes.

Electron multiplier photo-tubes, when applied to spectrographic analysis, avoid many of the complexities of the photographic method, but they may also introduce several new sources of error. In order to utilize these tubes successfully in spectrographic work, all tube characteristics and anomalies must be carefully considered.

One of the most serious limitations encountered in the use of multiplier photo-tubes is the presence in these tubes of a relatively high dark current. By dark current is meant that current which flows in the tube output when no light falls

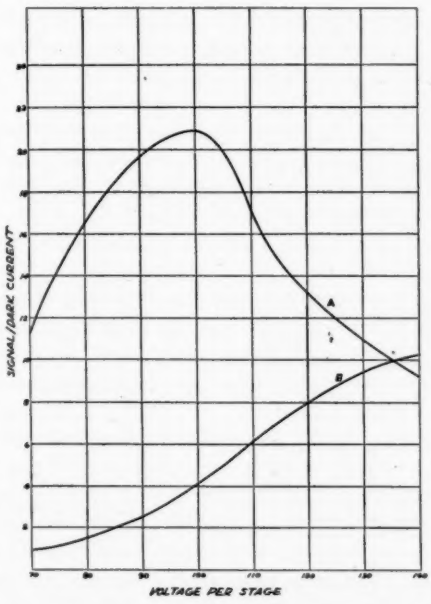

FIG. 1. Variation of the signal to dark current ratio as a function of the applied voltage per stage for a 931A tube (curve *A*) and a 1P28 tube (curve *B*).

[1] O. S. Duffendack and W. E. Morris, J. Opt. Soc. Am. **32**, 8 (1942).
[2] G. H. Dieke and H. M. Crosswhite, J. Opt. Soc. Am. **35**, 471 (1945).
[3] G. H. Dieke, H. Y. Loh, and H. M. Crosswhite, J. Opt. Soc. Am. **36**, 185 (1946).
[4] G. H. Dieke and H. M. Crosswhite, J. Opt. Soc. Am. **36**, 192 (1946).
[5] J. L. Saunderson, V. J. Caldecourt, and E. W. Peterson, J. Opt. Soc. Am. **35**, 681 (1945).

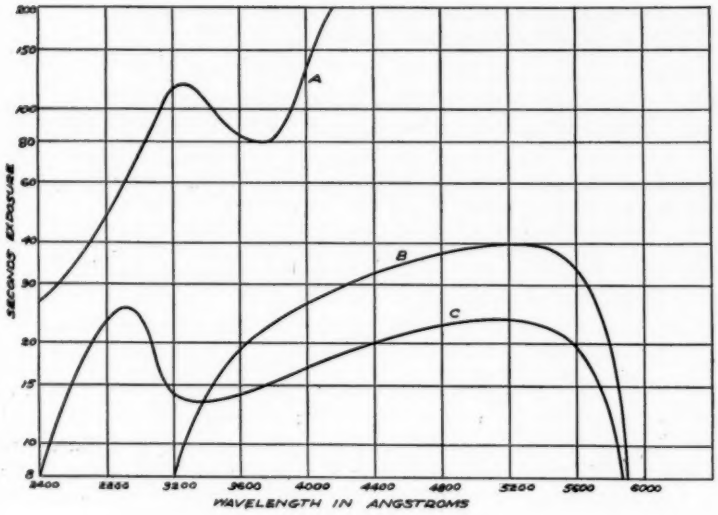

FIG. 2. Sensitivity of photo-tubes compared to sensitivity of photographic plates as a function of wave-length. Ordinate is in seconds of exposure necessary to produce a density of 50 percent on a photographic plate for an illumination level sufficiently high to yield a current ten times the dark current in the photo-tube. Curve *A* shows a comparison of the 1P28 with an S.A.-1 plate. Curves *B* and *C* are for a 931A and a 1P28, respectively, compared to a 103-B plate.

on the photo-cathode. The dark current present in multiplier photo-tubes is generally of the order of 10^{-7} to 10^{-8} ampere, varying greatly from tube to tube. Since this is of the same order of magnitude as the current produced by the light intensities usually encountered in spectrographic analysis, the presence of this dark current constitutes a serious limitation to their sensitivity and accuracy in quantitative applications.

The ratio of the photo-tube sensitivity to dark current as a function of the applied voltage per stage is given in Fig. 1. Since the dark current limits the ultimate sensitivity of the photo-tube, the ratio of sensitivity to dark current should be taken as the criterion for the selection of tubes and for the determination of optimum operating conditions. It is of interest to note that although tube No. 2 (curve *B*) has the higher sensitivity, its dark current is also higher, and the ratio of sensitivity to dark current is much lower than for tube No. 3 (curve *A*).

Since tube No. 3 showed a pronounced peak in its sensitivity to dark current ratio at about 100 volts, this value was selected as the optimum operating value for subsequent experiments. The curve for tube No. 2 does not have a peak, but shows a tendency to level off at about 140 volts. The use of voltages much greater than 100 volts per stage was considered inadvisable since the manufacturers recommend a maximum cathode-to-anode voltage of 1250 volts. One tube (No. 1) showed a strong tendency to spark at voltages above 120 volts per stage, and was also found to be generally erratic.

The current output of the photo-tubes as a function of the illumination has been checked for currents up to 30 microamperes and found to be linear within this range. Fatigue effects, as will be shown later, become very pronounced at high illumination levels, and it was found inadvisable to operate at levels above one microampere.

A comparison of the spectral response of a 931A and a 1P28 photo-tube with the response of several types of photographic emulsions is shown in Fig. 2. The abscissa is in angstroms, the ordinate is in seconds of exposure, to produce a density of 50 percent on the photographic plate when the illumination is at such a level as to produce a current of ten times the dark current in the photo-tube. The entrance slit on the photo-tube was 35 microns wide and 5 mm long.

From curve *C* we see that the sensitivity of the 1P28 photo-tube compares favorably with that of the 103-B emulsion from 2400 to 5800 angstroms. The 103-B type was chosen for this comparison because it has a very wide spectral response and a very high sensitivity. The spectrum analysis No. 1 plate (compared to the 1P28 tube in curve *A*) is the type most frequently used in industrial applications. Curve *B* shows that

the 931A tube cuts off at approximately 3400 angstroms.

Of the characteristics discussed above, only the high dark current presents a serious obstacle to the use of these tubes for the measurement of low light intensities such as are encountered in spectrographic analysis. There is, however, another phenomenon characteristic of these phototubes that seriously affects their usefulness to quantitative spectrographic analysis, namely, fatigue, i.e., the change of tube sensitivity with time. Any scheme suggested for the application of electron multiplier photo-tubes must take into account the limitations imposed by both dark current and tube fatigue.

Fatigue effect in a typical multiplier tube is shown in Fig. 3, with sensitivity plotted as a function of time. Curve *A* shows the fatigue resulting from illumination by a weak light yielding an initial current of approximately 3 microamperes. Curve *C* shows the fatigue effect due to a stronger light on the same tube. The initial current in this case was 23 microamperes. Fatigue curves of this type have been run for as long as twelve days without reaching an equilibrium value. From studies of such curves, it can be seen that the rate of fatigue is approximately proportional to the intensity of illumination, as pointed out by Dieke.[6]

Recovery from fatigue is shown in curve *B*, Fig. 3. This is a much slower phenomenon than fatigue itself. It is of interest to note what happens to the sensitivity of a tube if it is fatigued, allowed to recover partially, and then fatigued again. Curve *D* shows such a case. The conditions are the same as for curve *C* after the tube has been allowed to partially recover. It is apparent from studies of these fatigue curves that the rate of sensitivity change in a tube is dependent not only on the level of illumination and the initial sensitivity, but also on the previous history of the tube. It is important to remember this when considering the possibility of pre-fatiguing a tube with a strong light to attain equilibrium faster.

The direct current component of the dark current has been discussed above. It is also of interest to study the frequency components present in this dark current. Measurements of the dark current with a frequency analyzer showed a nearly constant signal over a constant frequency interval at all frequencies up to 17,000 cycles. This we would have expected, since the dark current is caused principally by spontaneous electron emission, together with the electron emission produced by the bombardment of the cathode surfaces by residual gas ions. Since these effects are random in nature, all frequencies are

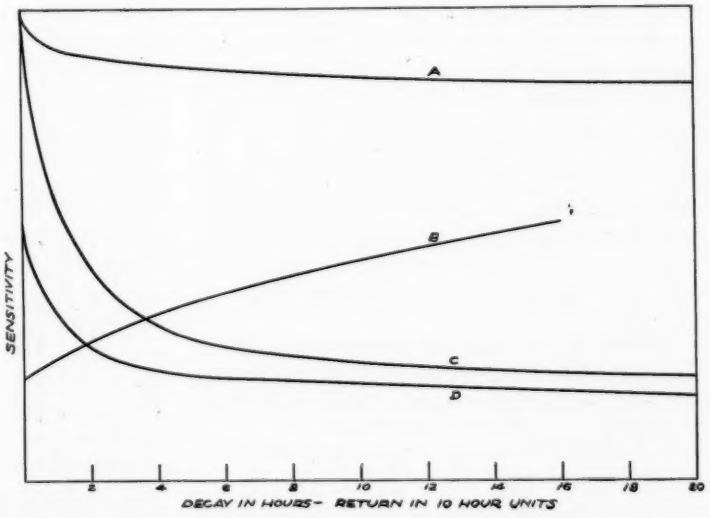

FIG. 3. Fatigue effect in a 931A photo-tube. Curve *A* is for low illumination, curves *C* and *D* for high illumination, and curve *B* shows return from a fatigued condition.

[6] G. H. Dieke and H. M. Crosswhite, J. Opt. Soc. Am. 35, 471 (1945).

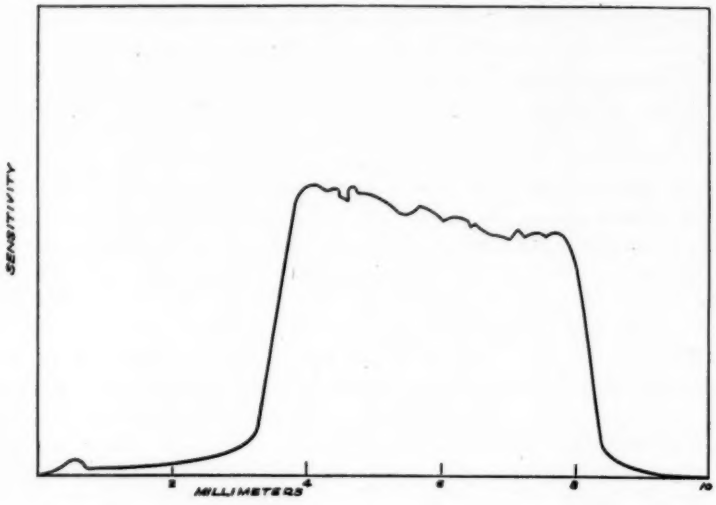

Fig. 4. Change in sensitivity from left to right (facing tube) across photo-surface of a 931A tube.

equally probable. The magnitude of the frequency component should conform to the familiar formula for shot effect:[7] In a single stage tube,

$$\bar{V}_f^2 = 2eR^2 I_0 \Delta f$$

where $\bar{V}_f$ is the r.m.s. voltage due to random noise over the frequency range between f and $(f+\Delta f)$. I_0 is the direct current component of the electron current, R is the resistor through which the current passes and across which the voltage $\bar{V}_f$ is measured, and e is the electronic charge.

In a nine stage photo-tube this formula must be modified slightly. Because of the multiplication by secondary emission at each dynode, one electron emitted from the photo-cathode will cause a packet of x^9 electrons to arrive at the plate. x is the mean electron multiplication per stage, and x^9 is the total multiplication of the tube. Similarly, one electron starting from the first dynode will result in a packet of x^8 electrons arriving at the plate. In general, the packet size will be x^n where n is the number of stages of amplification by secondary emission.

If we assume that the thermal emission is approximately the same for all stages, then each stage will contribute a fraction of the total dark current equal to

$$\frac{x^n}{x^9+x^8+\cdots+1}.$$

[7] W. Schottky, Ann. d. Physik **57**, 541 (1918).

If we apply these corrections to the shot effect formula we have

$$\bar{V}_f^2 = 2R^2\Delta f I_0 \sum_{n=0}^{9} (x^n e)\frac{x^n}{x^9+x^8+\cdots+1}$$

$$= 2R^2 e\Delta f I_0 \frac{x^{10}+1}{x+1}.$$

For a particular tube, the current amplification and the direct current component of the dark current were measured. The value of $\bar{V}_f$ calculated from the above formula for a frequency interval of 37 cycles was 2.38×10^{-4} volt through a resistance of 200,000 ohms. This is a reasonable check, within experimental error, on the measured value of 2.2×10^{-4} volt over the same frequency interval. This check would indicate that only a

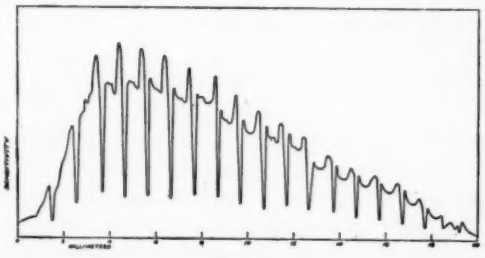

FIG. 5. Change in sensitivity from bottom to top of photo-surface of a 931A tube. The saw-tooth character of the curve is due to the grid wires of the tube.

negligible portion of the dark current in this particular tube was caused by ionic bombardment, since such a bombardment would liberate more than one electron on the average from the initial surface.

As has been suggested by Saunderson,[8] the dark current present in a tube may be readily cancelled out by means of a properly adjusted bucking voltage. Such a scheme will work only if the dark current remains constant with time. In the photo-tubes studied, two types of fluctuations were observed. The first had a period of the order of a second, with a maximum fluctuation of the order of ten percent about the mean value. The effect of such fluctuations can be minimized by taking readings that average over a time interval of ten seconds or more.

The second type of fluctuation noted was a general drift in the dark current. Over a period of minutes, this drift often changed the magnitude of the dark current by as much as 20 percent. Even greater changes were observed when a grounded wire was brought near the tube envelope, indicating that much of the fluctuation in the dark current may be due to the electrostatic influence of charges accumulated on the inner surface of the glass envelope of the photo-tube.

An attempt was therefore made to minimize this drift by wrapping the tube in a grounded metallic foil. When this foil was first placed on the tube and grounded, the dark current changed by as much as a factor of two, increasing in some tubes, decreasing in others. In each case, however, the current again slowly drifted back toward the value observed before grounding, and after a half-hour or more, the current again resumed its random drift.

The change of sensitivity over the photo-surface was studied by passing a narrow slit of light of constant intensity across the photo-surface. A vertical slit was first passed across the photo-surface from left to right, facing the tube. The curve obtained is shown in Fig. 4. The sensitivity is low over a large portion of the cathode, but has a fairly flat sensitive portion. The curve obtained by moving a horizontal slit from the bottom to the top of the photo-surface is shown in Fig. 5. The sensitivity here shows a sharp increase near the bottom of the tube and a slow decrease toward the top. The saw tooth character of the curve is caused by the shadow cast by the grid structure placed before the photo-cathode by the manufacturer as an electrostatic shield. From these curves we see that care must be taken to avoid even slight shifts in the image position on the photo-cathode.

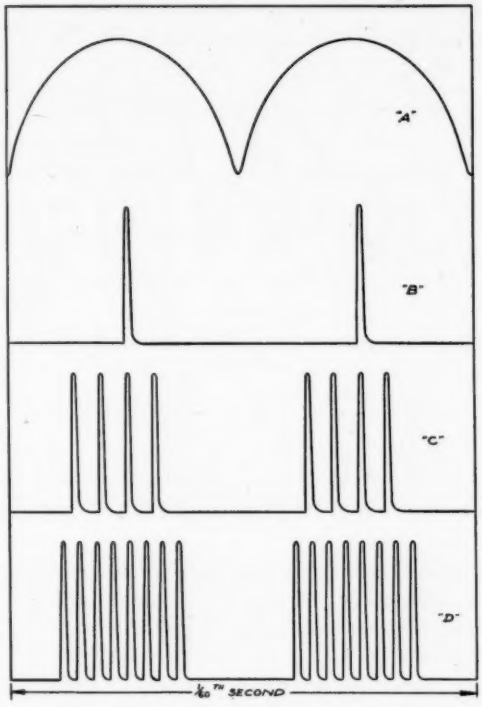

FIG. 6. Wave shapes of various light sources. Curve *A* is for an a.c. arc source, *B*, *C*, and *D* are for spark sources of 1, 4, and 8 sparks per half-cycle of supply voltage. The ordinate is light intensity, the range of the abscissa is 1/60 second.

SOURCE CHARACTERISTICS

Before attempting to design an instrument to be used in measuring spectral line intensities, it is advisable to consider the characteristics of the light sources used in spectrographic analysis.

In checking photo-tube characteristics, a low pressure mercury-argon discharge tube was used as a constant intensity light source. This type of tube has been checked photographically and

[8] J. L. Saunderson, V. J. Caldecourt, and E. W. Peterson, J. Opt. Soc. Am. **35**, 681 (1945).

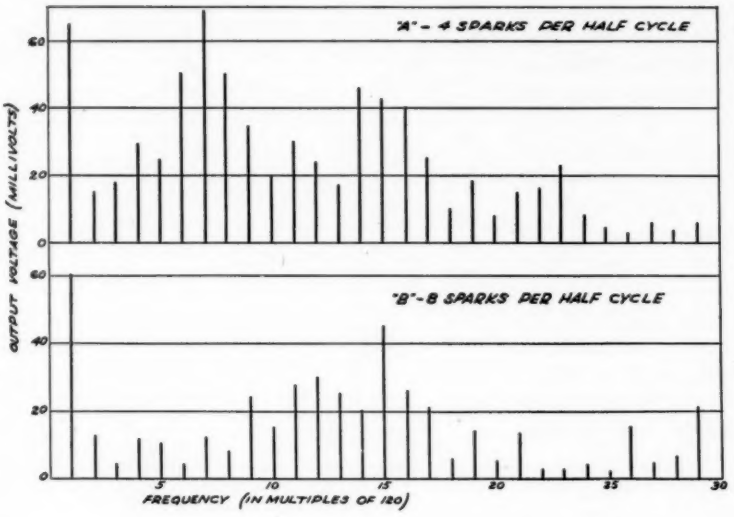

FIG. 7. Frequency components present in the light output of a spark source as measured by an electronic wave analyzer.

found to be constant within 2 percent over large time intervals. In routine spectrographic work, however, the most common types of sources are the a.c. arc and spark. The light emitted by an a.c. arc source is plotted as a function of time in curve *A*, Fig. 6. Since the frequency of the power supply was 60 cycles per second, the fundamental frequency of the light output is 120 cycles.

The spark source used is of the air-quenched interrupted spark type described by Fowler and Wolfe.[9] In this type of spark source, the number of spark discharges per half cycle of the supply voltage can be controlled by varying the applied transformer voltage. In curve *B*, Fig. 6, the light output is shown for a circuit condition of one spark per half-cycle. This excitation is similar to that produced by the Feussner type interrupted high voltage spark. Curves *C* and *D* show similar curves for four and eight sparks per half-cycle of supply voltage.

If the light output from these sources is to be converted into electrical impulses and amplified by any type of electronic circuit, it is important that we know what frequency components are predominant.

The a.c. arc light output resembles the output of a full wave rectifier, and by far the largest frequency component present is at 120 cycles. The spark source is more complex, and the frequencies present will depend to a large extent on the circuit constants used. Since the fundamental frequency of the spark source is 120 cycles for a symmetrical discharge, the output must contain only frequencies which are integral multiples of 120 cycles. If we know the wave shape of the light output of the spark discharge, it is possible to calculate the Fourier components of the light output. From these Fourier coefficients the relative amplitudes of the frequency components present in the light output of the spark source can be calculated.

The frequency components of the light output from the spark source can also be measured with an electronic wave analyzer connected to the photo-tube output. The frequency components measured in this manner are shown in Fig. 7 for several source conditions. Curve *A* shows the frequency components present in the light output from a spark discharge of 4 sparks per half-cycle. The same analysis for 8 sparks per half-cycle is shown in curve *B*. It was noted in studying a large number of source conditions that the 120-cycle component is always either the largest or one of the largest components present.

In comparing the intensity fluctuations of various types of sources, it was found that the arc source yields a fairly constant light intensity over a short period of several cycles. The arc source,

[9] R. G. Fowler and R. A. Wolfe, J. Opt. Soc. Am. **35**, 170 (1945).

however, shows great variations in intensity as the arc moves over the electrode surface. Not only the absolute intensity, but also the intensity ratios of spectrum lines undergo great changes as the arc moves from the molten metal to a fresh surface of solid metal or oxide.

The spark source, on the other hand, shows a large variation in intensity between individual sparks. Figure 6 shows an idealized picture of the spark source. In practice, the individual spark discharges are found to vary by as much as 50 percent between successive sparks. Over a long period average of several seconds, however, the spark source is found to be quite constant except for systematic variations which will be brought out later in a discussion of the intensity ratios of spectral lines.

CIRCUIT DESIGN REQUIREMENTS

In view of the photo-tube and spectrographic source characteristics outlined above, any instrument designed to measure the intensity ratios of spectral lines must meet certain qualifications listed below.

(1) The circuit must be sensitive enough to detect the very low levels of illumination encountered in spectrographic work. Except as a laboratory tool, high sensitivity galvanometers are impractical. Some type of electronic amplifier appears to be the only device suited to the requirements of routine analysis.

(2) The circuit used in conjunction with the electron multiplier photo-tubes should yield as high a signal to dark current ratio as possible. It is this ratio rather than the absolute tube sensitivity which usually determines the lower limit of intensities that can be measured. Most schemes proposed for the use of electron multiplier photo-tubes in spectrographic work have utilized the direct current component of the photo-tube output. The large dark current present can be partially eliminated by the use of a bucking voltage to neutralize the dark current. This scheme, however, does not compensate for the large fluctuations present in the dark current.

It is, however, possible further to improve the signal-to-noise ratio of the photo-tubes. We have shown above that the dark current output of the photo-tube has a continuous and uniform frequency spectrum, i.e., the output current within a fixed frequency interval is constant at all frequencies, and proportional to the width of the frequency interval used. The light output from the spark source as shown in Fig. 7, however, shows a discreet frequency spectrum. The energy here is concentrated in a finite number of narrow frequency bands.

A large increase in the signal-to-dark current ratio can be achieved by utilizing only that portion of the tube output which corresponds to one of the major frequencies present in the light output. The amount of dark current passed by a frequency selective device will be proportional to the band width detected by that device. If we take a frequency interval which includes only one of the frequency components of the source, the amount of signal passed by the detector will be independent of the band width of the detector until that band width becomes of the order of one cycle.

Since the 120-cycle component is either the predominant or at least one of the major frequency components of all sources studied, this component was chosen as the optimum frequency to be used in subsequent work.

(3) If 120-cycle signals are to be used in the vicinity of a high voltage spark source, the problem of electromagnetic pick-up must be carefully considered. In practice, the amount of interference was reduced to a satisfactory level by completely shielding all circuit elements and all connecting cables, by grounding the amplifier circuits, cable shields, and chassis at one point, and by placing a stage of preamplification in the same housing with each photo-tube.

(4) Any circuit used, in order to be readily applicable, must measure intensity ratios of spectral lines, and should measure the true ratio without being affected by variations in absolute intensity. If vacuum tube amplifiers are used, care must be taken not to overload the amplifiers.

(5) Instantaneous values of the intensity ratios of two spectral lines are usually very erratic, whereas an average taken over a short time interval will be found to be quite constant. The detector used must, therefore, be capable of measuring average, rather than instantaneous intensity ratios.

(6) To be useful in actual spectrographic analysis, the instrument used in measuring in-

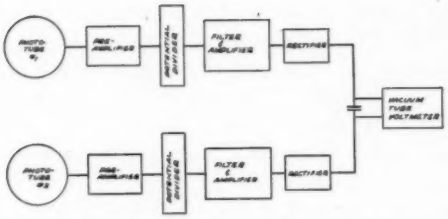

FIG. 8. Simplified box diagram of component parts of the ratio analyzer.

tensity ratios must be so designed that these ratios can be measured to an accuracy of 1 percent over the range of intensities to be used.

A CIRCUIT FOR MEASURING INTENSITY RATIOS

In most spectrographic analysis, the empirical intensity ratio of spectral lines is the quantity used to determine the composition of samples analyzed. Several methods of obtaining this ratio were tried, and the final circuit used is shown in simplified box diagram form in Fig. 8.

The power supply, shown in Fig. 12, furnishes 1263 volts d.c. to each photo-tube. A potential divider composed of ten 0.2 megohm resistors supplies the dynode voltages. This potential divider together with a preamplifier and an attenuator switch are included in a brass housing covering each photo-tube.

Since no suitable grating instrument was available, it was found necessary to use the photo-tubes on a Littrow instrument. In this type of spectrograph, the light strikes the focal plane at the exit slit at an angle of approximately 30°. It was therefore necessary to use a front surface aluminum mirror to direct the light from the entrance slit on the photo-tube housing to the photo-tube. The entrance slit is in the focal plane of the spectrograph. Light passing through the slit is reflected by the mirror onto the photo-surface of the tube. The mirror is so arranged that it can be pivoted readily. This adjustment allows the light to be directed to the most sensitive portions of the photo-surface.

After preamplification, the signal from each photo-tube passes from the photo-tube housing through a doubly shielded cable to a calibrated potential divider in the main cabinet. From each potential divider, the signal passes through a narrow band-pass filter adjusted to amplify only the 120-cycle component in the signal.

The 120-cycle components from each filter are again amplified and rectified. The difference between the rectified voltages from the two photo-tube circuits is applied across a condenser, and the voltage across the condenser is then measured with a vacuum tube voltmeter having a high input impedance and a very long time constant.

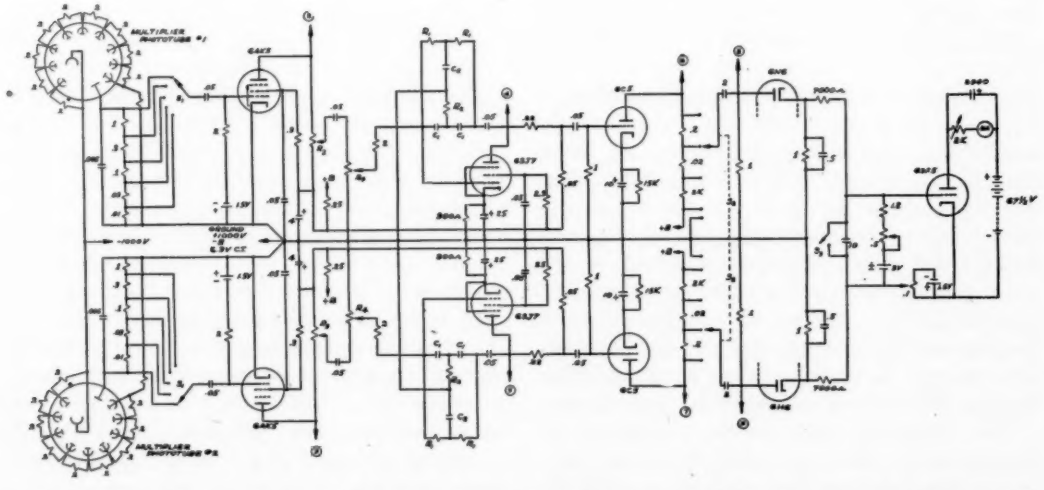

FIG. 9. Schematic wiring diagram of the ratio analyzer. Unless otherwise marked, all resistances are given in megohms, all capacitances in microfarads. $R_3=0.1$ meg.; $R_4=1.0$ meg.; $C_2=0.0016$ mf $=2C_1$; $R_1=2R_2=1/\omega C_1$ (adjusted for maximum gain where $\omega=240\pi$).

In operation, the method of analysis is essentially a null method. The calibrated potential dividers are adjusted to give equal output signals from the two amplifiers, as indicated by the vacuum tube voltmeter when the average condenser voltage is zero. At null balance, the ratio of the light intensities striking the two photo-tubes is proportional to the ratio recorded by the potential dividers.

A complete schematic diagram of the circuit is shown in Fig. 9. A condenser of 0.005 mf was inserted across the photo-tube load resistor. Since only the 120-cycle frequency components are amplified, this condenser is added to by-pass the higher frequency components. The five-position switch across the photo-tube output is used to decrease the signal amplitude for high intensity spectral lines. Since tube fatigue increases rapidly with increased signal amplitude, it was found advisable to use only positions 4 and 5 on this switch, and to cut down on the light intensity falling on the slit when overloading occured.

The type of potential dividers used is shown in Fig. 10. A calibrated variable resistor is used to interpolate between ten switch positions. Such an arrangement gives ample accuracy for all types of analyses over all but the lower portion of the potential divider. The first potential divider ($R_3=0.1$ megohm) comprises the plate resistor of the preamplifier tube. Its adjustment is only roughly calibrated since it is used only in adjusting the amplifier output so that the second potential divider may be operated over a convenient range.

The second pair of potential dividers ($R_4=1.0$ megohm) is carefully calibrated. In operation, this pair of potential dividers is adjusted to give a null balance on the vacuum tube voltmeter. The ratio of the resistance settings on the two potential dividers R_4 is then proportional to the ratio of the light intensities falling on the two photo-tubes. In the actual circuit, the two potential dividers shown as R_4 are not both variable. One of the two is variable whereas the other is merely a fixed resistor whose value is equal to the total resistance of the variable unit. The two resistors are then interchanged between the two circuits by means of a four pole double throw switch. By using this switch and one variable potential divider, all ratios are possible. To avoid complicating the diagram, this switch has not been inserted in Fig. 9.

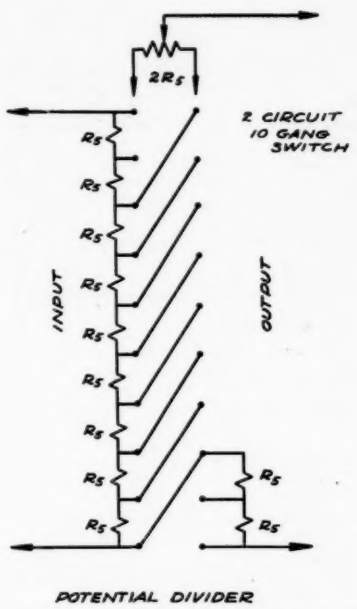

FIG. 10. Potential divider used for R_3 and R_4 in Fig. 9. For R_3, $R_s=0.01$ meg.; for R_4, $R_s=0.1$ meg.

From the potential dividers the signal passes through a narrow band-pass filter tuned to 120 cycles. The filter is an $R-C$ parallel T circuit on a pentode, so connected that the amplifier is degenerative at all frequencies except 120 cycles. The amplification of this type of filter when applied to a degenerative amplifier has a very sharp peak at resonance. The parallel T filter has been described by Scott[10] and more recently by Stanton.[11]

The actual response curve for the filter stage is plotted in Fig. 11. The filter constants were chosen with $C_1=0.00080$ mf, and the other constants to satisfy the following relations:

$$C_2=2C_1,$$
$$R_1=2R_2=1/\omega C_1,$$

where

$$\omega=2\pi f=240\pi.$$

Exact adjustment of the circuit constants was attained by applying a signal to the circuit from

[10] H. H. Scott, Proc. I.R.E. **26**, 226 (1938).
[11] L. Stanton, Proc. I.R.E. **34**, 447 (1946).

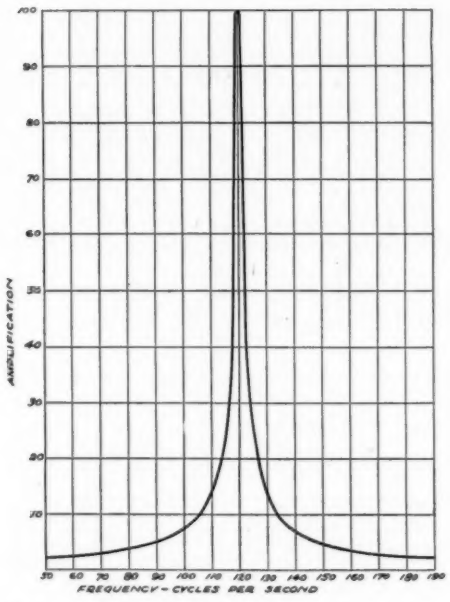

FIG. 11. Amplification of filter section as a function of frequency.

an electronic signal generator and adjusting the resistors to give a peak in the amplification at 120 cycles. The frequency adjustment can be made very accurately by using the filter output to form Lissajous figures with the power line frequency on an oscilloscope.

It is this filter which is responsible for most of the gain in signal to noise ratio achieved by the circuit. Since the filter is tuned to 120 cycles, one of the largest frequency components present in the source, a major portion of the signal amplitude is amplified by this stage. However, only that portion of the dark current which falls within the range of the filter response curve will be amplified. The resultant signal to noise ratio on the spark source using this filter is approximately eight to ten times as great as the ratio obtained without the filter.

After one more stage of amplification, the signal is rectified, and the voltage difference between the two circuits is applied across a condenser. The final stage of amplification is hardly necessary for voltage amplification, since the full output of the filter section is already sufficient to overload the final amplifier. This amplifier stage is utilized, however, to minimize coupling between the circuits leading from the two phototubes through the rectifier-condenser circuit.

The vacuum tube voltmeter used is designed to measure small changes in d.c. voltage across the condenser. It must have a high input impedance and a large time constant. The meter in the plate circuit of the vacuum tube voltmeter is a 1 ma full scale ammeter. In adjusting the circuit, the grid bias is set to give a reading of 0.5 ma in the plate circuit when switch S_3 is closed. With S_3 open and S_2 switched to a zero signal position, the resistance in the grid circuit is adjusted so that

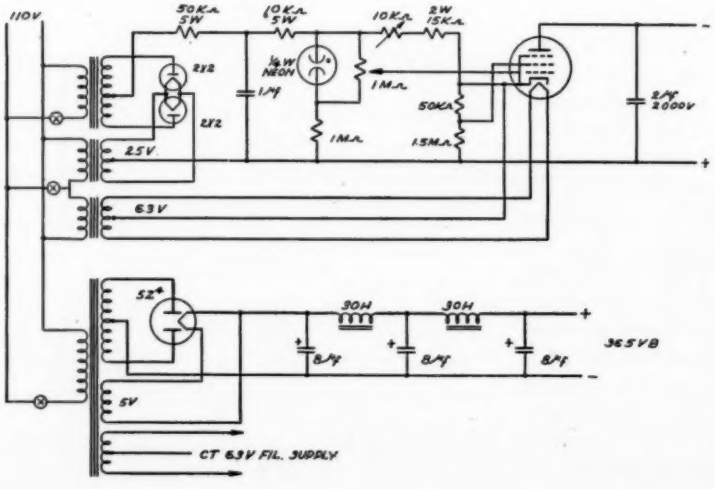

FIG. 12. Power supply for Fig. 9. The high voltage supply is 1263 volts d.c., the "*B*" supply is 365 volts.

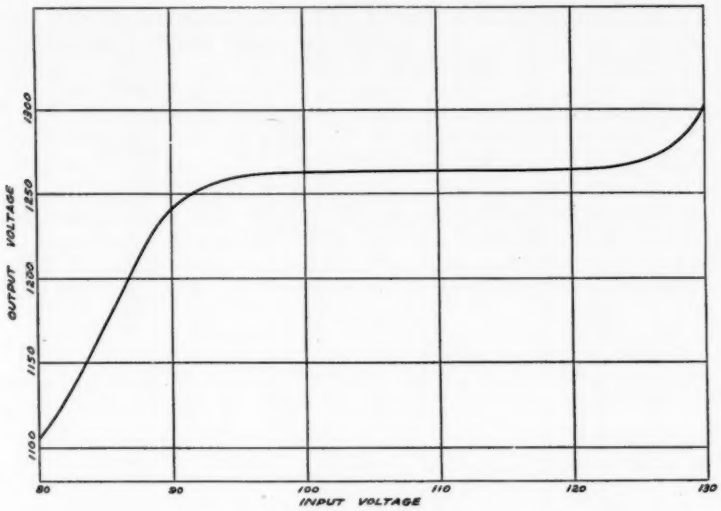

FIG. 13. Stability of output voltage of high voltage supply as a function of input voltage.

the ammeter again reads 0.5 ma. The three-volt battery in the grid circuit is necessary to balance out the voltage drop due to the flow of current through the resistors in the grid circuit.

The variable resistance in series with the milliammeter together with the 2000-mf condenser in parallel with the combination introduce a high variable time constant. This time constant can be adjusted to average over the random fluctuations in the intensity ratio.

The circled numbers on the circuit diagram indicate points where the oscillograph may be connected to the circuit. The connection is made through a rotary switch, and any point in the circuit can readily be checked. The oscillograph is essential as an indicating instrument when adjusting the photo-tubes to locate lines in the spectrum. It is also essential to use the oscillograph to check against overloading the amplifiers.

The power supply for the photo-tubes must be extremely stable. The tube output is very sensitive to changes in the power supply voltage. The power supplies used are shown in Fig. 12. The high voltage supply uses a pentode regulator. The stability of the high voltage output with changes in the input voltage is shown in Fig. 13. In practice, the power supply is used in conjunction with a commercial transformer type regulator. The ripple voltage in the output of the power supply is less than 0.02 percent.

APPLICATION TO QUANTITATIVE ANALYSIS

To test the apparatus under actual working conditions, the photo-tube circuit outlined above was used for the analysis of chromium in steel. The source used was the air quenched interrupted spark type with an inductance of 100 microhenries and a capacitance of 0.0045 mf. The source was operated at an interrupter breaking voltage of 14,000 volts with four sparks per half-cycle of the 60-cycle line frequency. The spectrograph used was a Hilger Littrow, slit width 35 microns, slit length 8 mm. The effective width of the photo-tube slit was 100 microns.

An analytical procedure was first tried using the 4254.34A chromium line. This line, however, proved quite unsatisfactory. The ratio of chromium to iron line intensity changed very rapidly as the pins became hot, and no satisfactory equilibrium was reached. Moreover, reproducibility of successive runs was poor. This same erratic behavior was also found photographically.

A chromium spark line was next selected, and much better results were obtained. The lines used were the 2677.17A chromium and the 2926.587A iron lines, both spark lines. Again, as in the case of the arc lines above, a change in intensity ratio was noted as the electrodes became hot. The change, however, was much less than previously, and reproducibility was within the error usually obtained in photographic procedures.

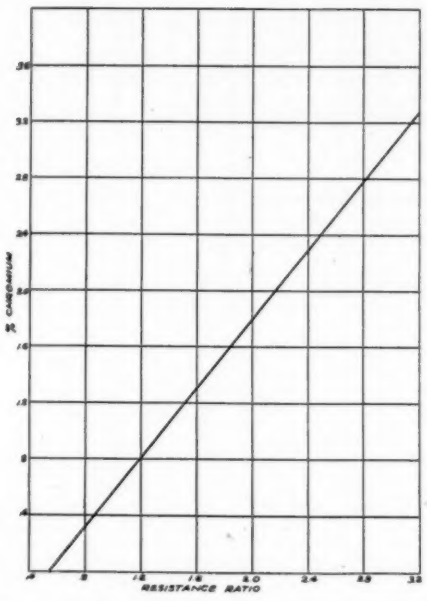

FIG. 14. Analytical curve for chromium analysis using photo-tubes on 2677.17A chromium and 2926.587A iron lines. Percent chromium plotted against ratio of resistances R_4 at null balance.

The analytical curve used for this analysis is shown in Fig. 14. Percent chromium is plotted against the resistor ratio at null balance. Reproducibility obtained for several alloys is shown in Table I, column 2.

Tube fatigue was found to be an important factor in obtaining reproducible results. A set of neutral filters was used in front of the spectrograph slit, and best results were obtained when the light intensity was such as to produce a current ten to twenty times as great as the dark current.

Reproducibility of results from day to day was also studied. In order to obtain satisfactory results, it was found necessary first to refocus the tubes on the lines used. Slight temperature changes in the room will shift the position of the line image on the exit slit. This phenomenon is caused both by the change in index of refraction of quartz and by the expansion of the metal parts of the spectrograph.

Since the sensitivity of the photo-tubes changes with time because of fatigue effects, it was also found necessary to reset the potential divider R_3 to compensate for any change in sensitivity. Best results were obtained when the tubes were first pre-fatigued for several minutes by exposing them to the light from the spark source. A standard sample of known analysis was then run. The potential divider R_4 was set at the value from Fig. 14 corresponding to the actual chromium content of the sample. The resistor R_3 was then adjusted for a null balance. Reproducibility obtained after making these adjustments is shown in Table I, column 4.

TABLE I. The analysis of chromium in steel.

Chemical analysis, %	Photo-tube analysis—I %	Error %	Photo-tube analysis—II %	Error %
2.95	2.94	0.3	2.98	1.0
2.95	2.97	0.7	2.96	0.3
2.95	2.88	2.4	2.94	0.3
2.95	2.91	1.3	2.89	2.0
2.95	2.98	1.0	2.97	0.6
2.95	2.98	1.0	2.90	1.7
2.00	2.05	2.5	2.03	1.5
2.00	2.01	0.5	1.98	1.0
2.00	1.98	1.0	2.02	1.0
2.00	1.98	1.0	2.02	1.0
2.00	1.99	0.5	1.97	1.5
2.00	2.02	1.0	2.00	0.0
1.40	1.40	0.0	1.42	1.4
1.40	1.38	1.4	1.40	0.0
1.40	1.41	0.7	1.37	2.1
1.40	1.36	2.8	1.39	0.7
1.40	1.44	2.8	1.41	0.7
1.40	1.39	0.7	1.42	1.4
0.82	0.83	1.2	0.84	2.5
0.82	0.79	3.6	0.81	1.2
0.82	0.81	1.2	0.81	1.2
0.82	0.84	2.5	0.78	4.9
0.82	0.82	0.0	0.79	3.6
0.82	0.84	2.5	0.79	3.6

As can be seen from Table I, the reproducibility of results obtained with electron multiplier phototubes compares favorably with that usually obtained through the photographic process, and the time required per analysis is less than two minutes. Studies of line intensities in the mercury-argon discharge tube also indicate that the errors found above cannot be attributed to the apparatus used, but, rather, are caused by fluctuations in the light output of the spark discharge. It was also noted that if the pins were allowed to become red hot, the intensity ratios became extremely erratic, again emphasizing the necessity for well selected and rigidly controlled source conditions.

Polarimetric Methods for the Determination of the Refractive Index and the Thickness of Thin Films on Glass

ANTONIN VAŠIČEK
The Technical University, Brno, Czechoslovakia

(Received May 6, 1946)

A general method, and various modifications thereof, are developed for the determination of the thickness and refractive index of thin films on glass from polarimetric measurements of the reflected light, using plane polarized incident light. The mathematical problem involved cannot be solved in explicit form but, as shown by numerical examples, exact results can be obtained by graphical interpolation.

IN a series of papers[1–4] the author has shown that the reflection of plane polarized light on a plane glass surface which is coated with a thin film is given by the relations:

$$\frac{R_p}{E_p}e^{i\delta_p}=\frac{r_p'+r_p''e^{-ix}}{1+r_p'r_p''e^{-ix}}, \qquad \frac{R_s}{E_s}e^{i\delta_s}=\frac{r_s'+r_s''e^{-ix}}{1+r_s'r_s''e^{-ix}}. \tag{1}$$

These relations take into account multiple reflections. $r_p=R_p/E_p$, $r_s=R_s/E_s$ are the ratios of the amplitudes of the reflected (R) and the incident (E) wave, the subscripts refer to components parallel (p) and normal (s) to the plane of incidence, and δ_p, δ_s are the phase changes of these components. r_p', r_s' and r_p'', r_s'' are the analogous amplitude ratios for single reflection at the surface of the film and the surface of the uncoated glass, respectively. They are given by Fresnel's equations

$$\begin{aligned} r_p'&=\frac{\tan(\varphi-j)}{\tan(\varphi+j)}, & r_s'&=-\frac{\sin(\varphi-j)}{\sin(\varphi+j)}, \\ r_p''&=\frac{\tan(j-r)}{\tan(j+r)}, & r_s''&=-\frac{\sin(j-r)}{\sin(j+r)}, \end{aligned} \tag{2}$$

where φ denotes the angle of incidence. j and r are the angles of refraction in the film and the glass, respectively. Hence, according to Snell's law

$$n=\frac{\sin\varphi}{\sin r}, \qquad n_0=\frac{\sin\varphi}{\sin j}, \tag{3}$$

where n is the refractive index of the glass and n_0 that of the film.

The quantity x in Eq. (1) is the optical path difference between the rays which are reflected from the air-film boundary and those reflected from the film-glass interface, i.e.,

$$x=\frac{2\pi}{\lambda}2n_0d\cos j. \tag{4}$$

λ is the wave-length of the light, d the thickness of the film. A film is called thin when $d<\lambda$.

[1] The reflection of light on glass with thin films. Rozpravy II. tr. Česká akademie věd a umění, Praha. No. 20, p. 37 (1941).
[2] "Critical remarks to Drude's theory of surface films," Česká akademie (1941).
[3] "Does compression influence the ellipticity of the reflected light?" Česká akademie (1942).
[4] "A new method for the measurements of strongly elliptically polarized light with half shade methods," Česká akademie (1943). All these papers are in Czech with a summary in German.

Equations (1) will be referred to as the generalized Fresnel's equations. In the previous papers[1] the exact validity of these relations has been demonstrated and it was shown[1] that they reduce in the limiting case of thin films to the equations given by Drude.[5]

The generalized Fresnel's equations (1) state that the light reflected from a coated surface is in general weakly elliptically polarized. They serve to compute the quantities accessible to optical measurements. These are the azimuth ψ of the restored plane of polarization, defined by

$$\tan\psi = r_p/r_s \tag{5}$$

and the phase difference $\Delta = \delta_p - \delta_s$. Taking the ratio of the two Eqs. (1) one obtains

$$\tan\psi e^{i\Delta} = \frac{\dfrac{r_p' + r_p'' e^{-ix}}{1 + r_p' r_p'' e^{-ix}}}{\dfrac{r_s' + r_s'' e^{-ix}}{1 + r_s' r_s'' e^{-ix}}} = \frac{\dfrac{r_p' + r_p'' \cos x - i r_p'' \sin x}{r + r_p' r_p'' \cos x - i r_p' r_p'' \sin x}}{\dfrac{r_s' + r_s'' \cos x - i r_s'' \sin x}{1 + r_s' r_s'' \cos x - i r_s' r_s'' \sin x}} = \frac{\dfrac{\alpha_1 - i\beta_1}{\alpha_2 - i\beta_2}}{\dfrac{\alpha_3 - i\beta_3}{\alpha_4 - i\beta_4}} \tag{6}$$

which is equivalent to

$$\tan\psi(\cos\Delta + i\sin\Delta) = \frac{(\alpha_1 - i\beta_1)(\alpha_2 + i\beta_2)(\alpha_3 + i\beta_3)(\alpha_4 - i\beta_4)}{(\alpha_2^2 + \beta_2^2)(\alpha_3^2 + \beta_3^2)}, \tag{7}$$

where

$$\begin{aligned} \alpha_1 &= r_p' + r_p' \cos x, & \beta_1 &= r_p'' \sin x, \\ \alpha_2 &= 1 + r_p' r_p'' \cos x, & \beta_2 &= r_p' r_p'' \sin x, \\ \alpha_3 &= r_s' + r_s'' \cos x, & \beta_3 &= r_s'' \sin x, \\ \alpha_4 &= 1 + r_s' r_s'' \cos x, & \beta_4 &= r_s' r_s'' \sin x. \end{aligned} \tag{8}$$

By equating real and imaginary parts on the two sides of Eq. (7) one is led to the following relations for the computation of ψ and Δ

$$\tan\psi\cos\Delta = \frac{(\alpha_1\alpha_2 + \beta_1\beta_2)(\alpha_3\alpha_4 + \beta_3\beta_4) - (\alpha_1\beta_2 - \alpha_2\beta_1)(\alpha_4\beta_3 - \alpha_3\beta_4)}{(\alpha_2^2 + \beta_2^2)(\alpha_3^2 + \beta_3^2)} = \frac{AA' - BB'}{(\alpha_2^2 + \beta_2^2)(\alpha_3^2 + \beta_3^2)}, \tag{9}$$

$$\tan\psi\sin\Delta = \frac{(\alpha_1\alpha_2 + \beta_1\beta_2)(\alpha_4\beta_3 - \alpha_3\beta_4) + (\alpha_3\alpha_4 + \alpha_3\beta_4)(\alpha_1\beta_2 - \alpha_2\beta_1)}{(\alpha_2^2 + \beta_2^2)(\alpha_3^2 + \beta_3^2)} = \frac{AB' + A'B}{(\alpha_2^2 + \beta_2^2)(\alpha_3^2 + \beta_3^2)}, \tag{10}$$

where

$$\begin{aligned} A &= \alpha_1\alpha_2 + \beta_1\beta_2 = r_p'(r + r_p''^2) + r_p''(r + r_p'^2)\cos x, \\ A' &= \alpha_3\alpha_4 + \beta_3\beta_4 = r_s'(1 + r_s''^2) + r_s''(1 + r_s'^2)\cos x, \\ B &= \alpha_1\beta_2 - \alpha_2\beta_1 = -r_p''(1 - r_p'^2)\sin x, \\ B' &= \alpha_4\beta_3 - \alpha_3\beta_4 = +r_s''(1 - r_s'^2)\sin x. \end{aligned} \tag{11}$$

Dividing Eq. (10) by Eq. (9) gives for the phase difference

$$\tan\Delta = (AB' + A'B)/(AA' - BB'). \tag{12}$$

---

[5] P. Drude, *Lehrbuch der Optik* (1912), third edition, p. 274.

To calculate the azimuth ψ we multiply Eq. (6) with its complex conjugate

$$\tan\psi e^{-i\Delta}=\frac{\dfrac{\alpha_1+i\beta_1}{\alpha_2+i\beta_2}}{\dfrac{\alpha_3+i\beta_3}{\alpha_4+i\beta_4}}. \tag{6'}$$

and obtain, by taking the square root of the product,

$$\tan\psi=\left(\frac{\dfrac{\alpha_1^2+\beta_1^2}{\alpha_2^2+\beta_2^2}}{\dfrac{\alpha_3^2+\beta_3^2}{\alpha_4^2+\beta_4^2}}\right)^{\frac{1}{2}}=\left(\frac{(\alpha_1^2+\beta_1^2)(\alpha_4^2+\beta_4^2)}{(\alpha_2^2+\beta_2^2)(\alpha_3^2+\beta_3^2)}\right)^{\frac{1}{2}}, \tag{13}$$

where

$$\begin{aligned}
\alpha_1^2+\beta_1^2&=r_p'^2+r_p''^2+2r_p'r_p''\cos x=(r_p'+r_p'')^2-4r_p'r_p''\sin^2\frac{x}{2},\\
\alpha_2^2+\beta_2^2&=1+r_p'^2r_p''^2+2r_p'r_p''\cos x=(1+r_p'r_p'')^2-4r_p'r_p''\sin^2\frac{x}{2},\\
\alpha_3^2+\beta_3^2&=r_s'^2+r_s''^2+2r_s'r_s''\cos x=(r_s'+r_s'')^2-4r_s'r_s''\sin^2\frac{x}{2},\\
\alpha_4^2+\beta_4^2&=1+r_s'^2r_s''^2+2r_s'r_s''\cos x=(1+r_s'r_s'')^2-4r_s'r_s''\sin^2\frac{x}{2}.
\end{aligned} \tag{14}$$

Equation (13) can be written in the form

$$\tan\psi=\left(\frac{(r_p'^2+r_p''^2+2r_p'r_p''\cos x)(1+r_s'^2r_s''^2+2r_s'r_s''\cos x)}{(1+r_p'^2r_p''^2+2r_p'r_p''\cos x)(r_s'^2+r_s''^2+2r_s'r_s''\cos x)}\right)^{\frac{1}{2}} \tag{15}$$

or, replacing $\cos x$ by $1-2\sin^2\frac{1}{2}x_1$, in the alternate form

$$\tan\psi=\left(\frac{\left[(r_p'+r_p'')^2-4r_p'r_p''\sin^2\dfrac{x}{2}\right]\left[(1+r_s'r_s'')^2-4r'r''\sin^2\dfrac{x}{2}\right]}{\left[(1+r_p'r_p'')^2-4r_p'r_p''\sin^2\dfrac{x}{2}\right]\left[(r_s'+r_s'')^2-4r_s'r_s''\sin^2\dfrac{x}{2}\right]}\right)^{\frac{1}{2}}. \tag{16}$$

Only the positive value of the square root has to be considered, because, by convention, $0\leqq\psi\leqq45°$.

The final result, represented by Eqs. (12) and (15) or (16), shows how the observable quantities ψ and Δ can be computed when the angle of incidence φ, the refractive index n of the glass, and the index n_0 and the thickness of the film are known. Thus for a film of known thickness and composition we can predict the state of polarization of the reflected light.

A GENERAL METHOD FOR THE DETERMINATION OF THE REFRACTIVE INDEX AND THE THICKNESS OF A THIN FILM

Of greater practical importance is the inverse problem, namely the determination of the thickness and index of a film on a glass of known index. Since polarimetric measurements furnish two quantities

it must, in principle, be possible to solve our two equation for the two unknowns n_0 and d. To be sure, the polarimeter used in our investigations, consisting of a Sénarmont compensator and a Nakamura[6] double plate, does not record the two quantities ψ and Δ which enter in the theory. It serves to determine the orientation of the major axis of the vibration ellipse of the reflected light and its ellipticity. The former is given by an angle χ with the plane of incidence, the latter by the ellipticity angle γ. But from these observed quantities we can readily compute the values of ψ and Δ according to the well known relations

$$\cos 2\psi = \cos 2\gamma \cos 2\chi, \quad \tan\Delta = \tan 2\gamma / \sin 2\chi.$$

In our measurements the incident light is always plane polarized at 45° or 135° to the plane of incidence.

The mathematical problem of finding the index n_0 and the thickness d of the film from ψ and Δ at a fixed angle of incidence involves the solution of the two Eqs. (12) and (15) or (16) with respect to n_0 and d. Unfortunately it is impossible to give these solutions in explicit form. However, we have succeeded in developing a method which furnishes the answer by means of a graphical interpolation.

The procedure consists in assuming first that the value of n_0 is approximately known within certain limits. We then chose three, or eventually more, values of n_0 within the range of these limits. For instance, we chose the values $n_0 = 1.350,\ 1.370,\ 1.390$ when the correct value is expected to be somewhere near 1.37. Using these fixed values of n_0 we next compute from Eq. (15) or (16) the values of the optical path difference x which would account for the observed value of ψ if the film would possess the chosen index. Each n_0 value will give a different x value. By entering these x values into Eq. (12) we then compute for each n_0 value the corresponding angle Δ. By plotting these Δ values *versus* n_0 we can then determine the correct value of n_0 which corresponds to the observed value of Δ. Having done this, it is then easy from the x *versus* n_0 curve to ascertain the correct value of x and, from Eq. (4), to evaluate the exact thickness of the film.

The general procedure outlined above contains only one step which needs further elaboration. This step deals with the computation of x corresponding to an assumed fixed value of n_0 and compatible with an observed value of ψ. There exist two alternatives to carry out this computation, either by using Eq. (15) or Eq. (16). Both alternatives involve the same procedure, namely, taking the square of either equation one gets

$$\tan^2\psi = \frac{(r_p'^2 + r_p''^2 + 2r_p'r_p'' \cos x)(1 + r_s'^2 r_s''^2 + 2r_s'r_s'' \cos x)}{(1 + r_p'^2 r_p''^2 + 2r_p'r_p'' \cos x)(r_s'^2 + r_s''^2 + 2r_s'r_s'' \cos x)}. \tag{15'}$$

or

$$\tan^2\psi = \frac{\left[(r_p' + r_p'')^2 - 4r_p'r_p'' \sin^2\frac{x}{2}\right]\left[(1 + r_s'r_s'')^2 - 4r_s'r_s'' \sin^2\frac{x}{2}\right]}{\left[(1 + r_p'r_p'')^2 - 4r_p'r_p'' \sin^2\frac{x}{2}\right]\left[(r_s' + r_s'')^2 - 4r_s'r_s'' \sin^2\frac{x}{2}\right]}. \tag{16'}$$

By substituting into the first equation

$$\begin{aligned} & f = r_p'^2 + r_p''^2, \quad && f' = r_s'^2 + r_s''^2, \\ \xi = \cos x, \quad & g = 1 + r_p'^2 r_p''^2, && g' = 1 + r_s'^2 r_s''^2, \\ & h = 2r_p'r_p'', && h' = 2r_s'r_s''; \end{aligned} \tag{17}$$

[6] A description of our polarimetric instruments and methods is presented in the author's paper, Phys. Rev. **57**, 925 (1940).

and by substituting into the second equation

$$\xi=\sin^2\frac{x}{2},\quad \begin{aligned} f&=(r_p'+r_p'')^2, & f'&=(r_s'+r_s'')^2,\\ g&=(1+r_p'r_p'')^2, & g'&=(1+r_s'r_s'')^2,\\ h&=-4r_p'r_p'', & h'&=-4r_s'r_s'' \end{aligned} \tag{17'}$$

one finds that both can be written in identical form, namely

$$\tan^2\psi=(f+h\xi)(g'+h'\xi)/(g+h\xi)(f'+h'\xi). \tag{18}$$

or

$$\xi^2[hh'(1-\tan^2\psi)]+\xi[fh'+g'h-\tan^2\psi(f'h+gh')]+fg'-f'g\tan^2\psi=0. \tag{19}$$

This is a quadratic equation for ξ, with the solution

$$\xi=-b_{(-)}^{+}(b^2-4ac)^{\frac{1}{2}}/2a, \qquad (a>0) \tag{20}$$

where the following abbreviations have been introduced

$$a=hh'(1-\tan^2\psi),\ b=fh'+g'h-\tan^2\psi(f'h+gh'),\ c=fg'-f'g\tan^2\psi. \tag{21}$$

Equation (20) solves the problem of finding x in terms of known or given quantities. The solution is unique, because only the positive sign of the square root must be considered. However, since ξ stands for $\cos x$ or $\sin^2\frac{1}{2}x$, respectively, it is necessary to decide in which quadrant the angle x must be placed. The answer depends on whether the film has a higher or lower index than the glass. If n_0 is smaller than n and if we chose the first alternative, Eq. (15), the quadrant is determined by the following table:

If the measured ellipticity is	and the calculated $\cos x$ is	then x lies in the quadrant
positive ($\gamma>0$)	positive	I or V
positive	negative	II
negative ($\gamma<0$)	negative	III
negative	positive	IV

The limiting cases, $\cos x=\pm 1$ correspond to maximum (+) and to minimum reflection. When n_0 is larger than n the above table is changed only insofar as the words positive and negative must be exchanged in the first column.

The decision of the quadrant is easier when the second alternative is employed. When $n_0<n$ the angle $\frac{1}{2}x$ is in the I quadrant when the ellipticity is positive, and it is in the II quadrant when γ is negative. For $n_0>n$ this statement is inverted. The maximum of reflection corresponds to $\sin^2\frac{1}{2}x=0$, the minimum to $\sin^2\frac{1}{2}x=1$.

With x placed in the correct quadrant it is then necessary only to evaluate the values of A, A', B, B', defined by Eq. (11) and to introduce them into (12) to complete the general procedure outlined above.

A special case is encountered when $r_p'=0$, which occurs for an angle of incidence of about 55°. For this case $h=0$ and the quadratic equation is reduced to a linear equation

$$\xi(fh'-gh'\tan^2\psi)=-(fg'-f'g\tan^2\psi), \tag{22}$$

with the solution

$$\xi=-fg'+f'g\tan^2\psi/h'(f-g\tan^2\psi). \tag{23}$$

The type of calculations involved in this procedure, and the accuracy of the data obtained, is illustrated by a numerical example in Table I and the corresponding diagram for the interpolation in Fig. 1. The measurements refer to an angle of incidence of 70° and were made with Na light

$\lambda = 5893$A. The film was expected to possess an index between 1.35 and 1.39 and the computations were made with the values 1.3500, 1.3700, 1.3900. The Δ *versus* n_0 graph in Fig. 1 shows that to account for the measured phase difference of 11°52.5′ a refractive index of $n_0 = 1.3632$ must be assigned to the film, and from the d *versus* n_0 curve it follows that its thickness is $d = 1.170\ 10^{-4}$ millimeter. It is clear that still higher precision could be attained by calculating additional points on the interpolation curves.

DETERMINATION OF THICKNESS AND INDEX OF FILMS FROM THE PRINCIPAL ANGLE OF INCIDENCE

In the preceding discussion the angle of incidence is arbitrary and the general method is applicable to any angle φ. It is therefore possible to choose certain preferred angles of incidence, for which the computation becomes simplified. One such angle is the principal angle of incidence, which we denote with φ_0. All other quantities relating to the case of principal incidence will carry the subscript 0. Principal incidence is characterized by the fact that one of the principal axes of the vibration ellipse of the reflected light is in the plane of incidence. This implies that χ is either 0° or 90°, Δ is either 90° or 270°, and $\psi_0 = \gamma_0$. The measurements at principal incidence furnish therefore the two angles φ_0 and ψ_0. Both angles are obtained by interpolation of a series of measurements in the neighborhood of principal incidence.

The mathematical condition for principal incidence is

$$A_0 A_0' - B_0 B_0' = 0. \tag{24}$$

Using the definitions (11) and substituting

$$\begin{aligned} &r_{p,0}'(1+r_{p,0}''^2) = p_0, \quad r_{p,0}''(1+r_{p,0}'^2) = q_0, \quad -r_{p,0}''(1-r_{p,0}'^2) = s_0, \\ &r_{s,0}'(1+r_{s,0}''^2) = p_0', \quad r_{s,0}''(1+r_{s,0}'^2) = q_0', \quad r_{s,0}''(1-r_{s,0}'^2) = s_0'; \end{aligned} \tag{25}$$

the condition (24) reduces to

$$(p_0 + q_0 \cos x_0)(p_0' + q_0' \cos x_0) - s_0 s_0' \sin^2 x_0 = 0, \tag{26}$$

which represents a quadratic equation for $\cos x_0$ with the solution

$$\begin{aligned} &(q_0 q_0' + s_0 s_0') \cos^2 x_0 + (p_0' q_0 + p_0 q_0') \cos x_0 + (p_0 p_0' - s_0 s_0') = 0, \\ &\cos x_0 = \frac{-(p_0' q_0 + p_0 q_0')_{(-)}^{+} [(p_0' q_0 + p_0 q_0')^2 - 4(q_0 q_0' - s_0 s_0')(p_0 p_0' - s_0 s_0')]^{\frac{1}{2}}}{2(q_0 q_0' + s_0 s_0')}. \end{aligned} \tag{27}$$

Taking $q_0 q_0' - s_0 s_0' > 0$, only the positive sign of the root has a physical meaning. The quadrant of x_0 is determined by the same table as used in the general method, the sign of $\cos x_0$ is the same as that of $(s_0 s_0' - p_0 p_0')$.

The computation of index and thickness of the film proceeds in this case as follows: Choosing a set of fixed values for n_0 the corresponding values of x_0 are computed from (27). The results are now introduced into Eqs. (15) or (16) and a graph of ψ *versus* n_0 is constructed, from which we can read the actual value of n_0, corresponding to the observed angle ψ, and the actual thickness of the film can be deduced from the x_0 *versus* n_0 plot.

This procedure involves a smaller amount of numerical computation than the general method; on the other hand it requires a larger amount of experimental data to permit accurate interpolation for the determination of the principal angle of incidence. The accuracy of the results is as good as by any of the other methods. The best degree of accuracy is obtained when the principal angle of incidence is close to the angle for minimum reflection. By using this method the author has been able to measure the dispersion of the film. A numerical example of the results obtained by this method is given in Table II.

DETERMINATION OF INDEX AND THICKNESS OF FILMS FROM THE ANGLE OF INCIDENCE WHICH GIVES A MINIMUM OF REFLECTION

There exist two other sets of angles of incidence for which the theory takes a simplified form. One set corresponds to the angles for which the reflected intensity reaches a maximum. The author has shown previously[6] that the state of polarization of the reflected light for maximum reflection is independent of the film, i.e., the reflected light is the same as if the film were not on the glass. Hence these angles cannot be used for our purpose. However, the angles of incidence for which the reflected intensity is a minimum do depend on the film and offer a possibility for the determination of its index and thickness.

TABLE I. The calculation of the refractive index, n_0, and of the thickness, d, of a thin interference film on glass by the general method.

$n=1.5687$ $\lambda=5893\text{Å}$

Measured $\varphi=70°00.0'$, $\psi=29°04.0'$, $\Delta=+11°52.5'$

n_0	1.3500	1.3700	1.3900
$\sin r$	0.59903	0.59903	0.59903
r	36°48′04″	36°48′04″	36°48′04″
$\sin j$	0.69608	0.68592	0.67605
j	44°06′47″	43°18′28″	42°32′09″
$\varphi-j$	25°53′13″	26°41′32″	27°27′51″
$\varphi+j$	114°06′47″	113°18′28″	112°32′09″
$j-r$	7°18′43″	6°30′24″	5°44′05″
$j+r$	80°54′51″	80°06′32″	79°20′13″
r_p'	−0.21723	−0.21661	−0.21568
r_s'	−0.47833	−0.48910	−0.49932
r_p''	0.020520	0.019887	0.018909
r_s''	−0.12889	−0.11503	−0.10168
f	0.047603	0.047315	0.046874
g	1.0000199	1.0000186	1.0000166
h	−0.0089146	−0.0086154	−0.0081562
f'	0.24541	0.25245	0.25966
g'	1.00380	1.00317	1.00258
h'	0.12330	0.11252	0.10154
ψ	29°04′00″	29°04′00″	29°04′00″
$\tan\psi$	0.55583	0.55583	0.55583
$\tan^2\psi$	0.30894	0.30894	0.30894
$1-\tan^2\psi$	0.69106	0.69106	0.69106
a	−0.00075962	−0.00066993	−0.00057234
b	−0.040498	−0.037408	−0.034134
c	−0.028035	−0.030530	−0.033226
$\cos x$	−0.70168	−0.82770	−0.98980
x	134°33′41″	145°51′45″	171°48′30″
$\sin x$	0.71250	0.56119	0.14249
A	−0.23239	−0.23393	−0.23534
A'	−0.37515	−0.37760	−0.37876
B	−0.013931	−0.010637	−0.0025689
B'	−0.070820	−0.049110	−0.010876
$\tan\Delta$	0.25156	0.17656	0.039641
Δ	14°07.0′	10°01.0′	2°16.0′
$\cos j$	0.71798	0.72768	0.73685
d mm	$1.1362\cdot10^{-4}$	$1.1976\cdot10^{-4}$	$1.3666\cdot10^{4}$

By the graphical interpolation (Fig. 1) $n_0=1.3632$, $d=1.170\cdot10^{-4}$ mm.

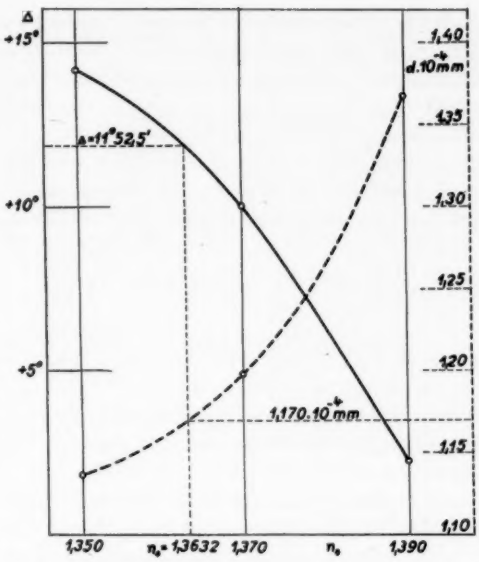

FIG. 1. Graphical interpolation for the determination of the refractive index and the thickness of a thin film on glass.

For minimum reflection the reflected light is strictly plane polarized, $\gamma=0$, and the polarimetric measurements furnish therefore only one angle, namely the azimuth $\psi_{\min}$, which in accordance with our conventions is always considered a positive angle. Denoting all quantities pertaining to this case with a subscript "min" we obtain the necessary equations from the condition that for minimum reflection the optical path difference is an odd multiple of $\frac{1}{2}\lambda$. Thus for the first minimum

$$2n_0 d\cos j=\tfrac{1}{2}\lambda.$$

The generalized Fresnel relations (1) are then

$$(r_p)_{\min}=\frac{r_p'-r_p''}{1-r_p'r_p''}=\frac{\dfrac{\tan(\varphi-j)}{\tan(\varphi+j)}-\dfrac{\tan(j-r)}{\tan(j+r)}}{1-\dfrac{\tan(\varphi-j)}{\tan(\varphi+j)}\dfrac{\tan(j-r)}{\tan(j+r)}},$$

$$(r_s)_{\min}=\frac{r_s'-r_s''}{1-r_s'r_s''}=-\frac{\dfrac{\sin(\varphi-j)}{\sin(\varphi+j)}-\dfrac{\sin(j-r)}{\sin(j+r)}}{1-\dfrac{\sin(\varphi-j)}{\sin(\varphi+j)}\dfrac{\sin(j-r)}{\sin(j+r)}} \tag{28}$$

TABLE II. The calculation of the refractive index, n_0, and of the thickness, d, of a thin interference film on glass from the principal angle of incidence.

$n=1.7393$	Measured $\varphi_0=52°20.0'$, $\gamma_0=+7°56.0'$			$\lambda=5893$A
n_0	1.4650	1.4700	1.4750	1.4800
γ_0	$+8°30.0'$	$+8°06.5'$	$+7°42.5'$	$+7°18.0'$
d mm	$9.083\cdot10^{-5}$	$9.114\cdot10^{-5}$	$9.154\cdot10^{-5}$	$9.204\cdot10^{-5}$

By the graphical interpolation $n_0=1.4721$, $d=9.131\cdot10^{-5}$ mm.

and furnish

$$\tan\psi_{\min}=\tan(\chi_{\min}-90°)=\frac{(r_p)_{\min}}{(r_s)_{\min}}. \qquad (29)$$

In this case it is more convenient not to use n_0 but the angle of refraction j of the film as parameter. Using the observed value of $\varphi_{\min}$ several values of $\psi_{\min}$, corresponding to at least three different j values are computed from Eq. (29). The correct values of j, n_0, and d are again ascertained by the interpolation procedure. This method again requires less computations than the general method but it involves more experimental work, since the position for minimum reflection must be determined by interpolation, i.e., one must determine the value of φ for which the ellipticity passes from positive to negative values. An example of the results obtained is given in Table III.

DETERMINATION OF THE REFRACTIVE INDEX OF FILMS WHEN THE PRINCIPAL ANGLE OF INCIDENCE PRODUCES MINIMUM REFLECTION

By careful control it is possible to produce films for which the principal angle of incidence coincides with the angle for minimum reflection. While it is difficult to satisfy this experimental condition, the evaluation of the data is exceedingly simple for this case. Since for this type of reflection γ and ψ become zero, we get $r_p'-r_p''=0$ and the generalized Fresnel equations furnish the trigonometric equation

$$\frac{\tan(\varphi-j)}{\tan(\varphi+j)}-\frac{\tan(j-r)}{\tan(j+r)}=0, \qquad (30)$$

which can be solved without resorting to the process of interpolation. The solution of (30) is

$$\sin 2j=(\sin 2\varphi\cdot\sin 2r)^{\frac{1}{2}}. \qquad (31)$$

It is then only necessary to introduce the j value into Snell's law to arrive at the refractive index of the film

$$n_0=\sin\varphi/\sin j.$$

DETERMINATION OF THE INDEX OF FILMS FROM TWO MINIMA OR MAXIMA FOR TWO WAVE-LENGTHS

The principle of this method has been discussed in a previous paper,[7] in which only the reflection maxima were considered. The relation which was derived for the maxima holds also for the angles of incidence $\varphi_{\min}$ for minimum transmission. Measurements of the 1st minimum for two different wave-lengths λ and λ' give

$$n_0=\left(\frac{\lambda^2\sin^2\varphi_{\min}'-\lambda'^2\sin^2\varphi_{\min}}{\lambda^2-\lambda'^2}\right)^{\frac{1}{2}}. \qquad (32)$$

While the determination of n_0 by this equation is the most convenient and experimentally most simple method it involves relatively large errors and is less accurate than the other polarimetric methods.

CONCLUDING REMARKS

The polarimetric measurements discussed in this paper presume that an approximate value of the index for the film is known. In particular, to assign the angle x to the correct quadrant it must be known whether n_0 is smaller or larger than n. For a film of unknown composition the following procedure permits a simple decision: First determine Brewster's angle $\bar{\varphi}$ for complete polarization of natural light, then measure very

TABLE III. The calculation of the refractive index, n_0, and of the thickness, d, of a thin interference film on glass from the minimum of the reflection.

$n=1.7393$	Measured $\varphi_{\min}=48°36.0'$, $\chi-90°=-1°26.5'$			$\lambda=5893$A
n_0	1.4700	1.4750	1.4800	1.4850
$\chi_{\min}-90°$	$-0°18.5'$	$-0°52.5'$	$-1°26.5'$	$-1°58.0'$

By the interpolation $n_0=1.4800$.
From the condition for the I minimum $2n_0d\cos j=d/2$ is $d=1.155\cdot10^{-4}$ mm.

[7] A. Vašiček, Phys. Rev. **57**, 925 (1940). The values of d in column 5 of Table I and in column 6 of Table II of this paper are incorrect. The given values should be multiplied by a factor 2.

roughly the principle angle of incidence. If $\varphi_0 < \bar{\varphi}$, then $n_0 < n$ and *vice versa* $\varphi_0 > \bar{\varphi}$ implies $n_0 > n$. The discussion of the method involving the angle of incidence for minimum transmission applies only to the case when $n_0 < n$. For $n_0 > n$ the same discussion is still valid provided the word minimum is everywhere replaced by maximum, and *vice versa*.

The evaluation of the optical path difference x by any one of our methods is indeterminate insofar as it cannot distinguish x values differing by a multiple of 2π. For thick films this indeterminacy can readily be removed by repeating the measurements with light of different wavelength.

In deciding which one of the polarimetric methods should be applied our experience has shown that it is best to attempt first to use the angle of incidence which gives minimum reflection for a film with $n_0 < n$, or a maximum for a film with $n_0 > n$. This method incurrs the least amount of calculations. If the method is found to be deficient we next attempt to measure the principal angle of incidence and only if this method does not give the desired accuracy do we resort to the general method, which involves the most amount of numerical work. The general method is best applied to a large angle of incidence between 50° to 70°.

The work entailed in these calculations could be greatly reduced by a set of tables for the pertinent quantities. Should the polarimetric method be adopted for routine tests of coated surfaces it would be advisable to construct such tables for various glasses and for films of various indices. These tables could be constructed for a fixed angle of incidence, but they must cover a range of angles ψ. Such tables would permit a very rapid determination of index and thickness of any film on glass.

Spectrochemical Analysis of High Copper in Cast Iron and Steel

W. R. KENNEDY
American Cast Iron Pipe Company, Birmingham, Alabama

(Received October 1, 1946)

Analysis of copper in steel and cast iron is easily carried out spectrographically with spark excitation in the range .02 to .20 percent using the 3274A copper line and the 3264A iron line, and in the range .20 to .80 percent with the 3286A iron line. These methods, described in detail herein, are not reliable for higher percentages of copper in cast iron and steel due to reversing of the 3274A copper line and segregation of copper in the iron with increasing copper concentrations. A unique method is described in detail which enables the spectrographer to analyze copper in cast iron and steel up to 2.50 percent (higher if standards are available). The method advocates the use of the a.c. arc at 5 kv and 1½ amperes for two reasons. The analysis lines chosen for calibration using the general internal standard method are the 2824A copper line and the 2819A iron line. Since copper percentages do not normally exceed .80 percent in the steel to be spectrographically analyzed by our laboratory, the technique described herein was developed primarily for cast iron in which copper percentages of three percent are not unusual.

FOR many applications copper has been found to be a useful and economical alloy in iron and steel.[1] Accurate analysis of copper in the higher percentages has been a source of much concern to the spectroscopist for two main reasons. In the first place, copper in iron has a tendency to segregate from solid solution as submicroscopic particles when the content is relatively high, so that results are extremely erratic. Furthermore, there are only two copper lines in the region used which are suitable for calibration purposes in the routine analysis of this element with the spark.

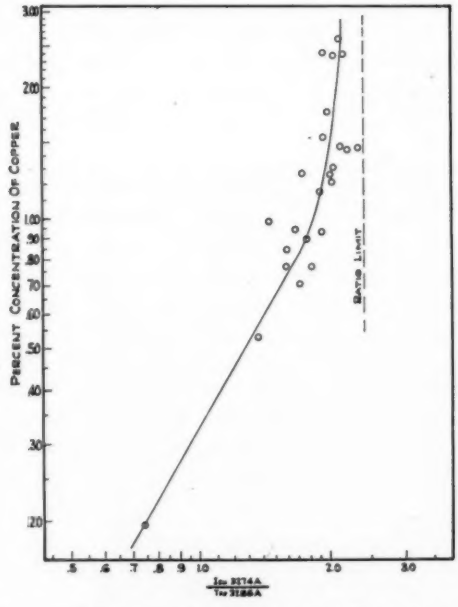

FIG. 1. Showing reversal of 3274A copper line using spark excitation.

Experience with copper analysis proves that the 3274A copper line is conveniently calibrated with spark excitation in both steel and cast iron in the range 0.02 to 0.20 percent with the 3264A Fe, and in the range 0.20 to 0.80 percent with the 3286A Fe. Details of this procedure are given in Table I and under General Procedure.

TABLE I. Details of procedures. All samples photographed through Corning No. 9800 glass filter over a 75-micron slit on Eastman Spectrum Analysis No. 1 35-mm film in the second-order range 2700 to 3400A with an analytical gap of 2 mm.

Percent range of copper	0.02 to 0.20 percent	0.20 to 0.80 percent	0.50 to 2.50 percent
Line pair used	Cu 3274A / Fe 3264A	Cu 3274A / Fe 3286A	Cu 2824A / Fe 2819A
Source	Spark		Arc
Sample-Steel	6-mm diameter pins unchilled; machined to 120° angle		
Cast iron	$\frac{7}{32}$-in. chilled pins; ground to 120° angle		
Preburn	10 seconds		30 seconds
Exposure	about 13 seconds*		16 seconds
Capacity (μf)	0.007		——
Inductance (mh)	0.045		——
Setting on ARL unit	¾ kva; L-1		5 kv; 1½A
Rotating sector	⅔ transmission		No sector
Air across analytical gap	¼ p.s.i.		No air

* Photoelectrically controlled.

[1] G. K. Manning and P. C. Rosenthal, Mining and Metall. 26, 601–605 (1945).

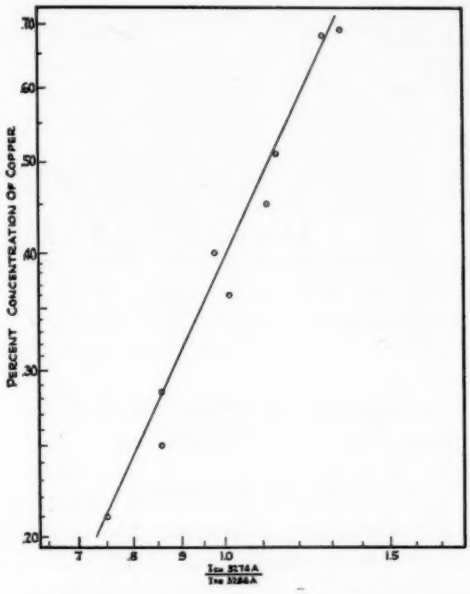

FIG. 2. Copper in steel—0.15–0.80 percent.

Because of the reversal of the 3274A Cu, its use above concentrations of 0.80 percent is unwarranted. For an example of the reversal effects on the ratio ICu 3274/IFe 3286, see Fig. 1. Such a curve cannot be used for calibration over the entire range, since a saturation point is reached and the intensity ratio does not increase further with increased copper content. Figure 1 also shows that reproducibility is poor in the high concentrations.

Failure to reproduce results on high copper contents with spark excitation prompted inspection of the a.c. arc as a means of making the sample more homogeneous and of permitting the use of a copper line not affected by reversal. The results obtained with the a.c. arc are considerably more precise and reproducible than with the spark.

Since the copper content in steel does not normally exceed .80 percent in this foundry, it was not necessary to subject steel to arc analysis. In cast iron, however, copper content up to 3 percent is not unusual, so the arc method was developed primarily for copper alloyed cast iron.

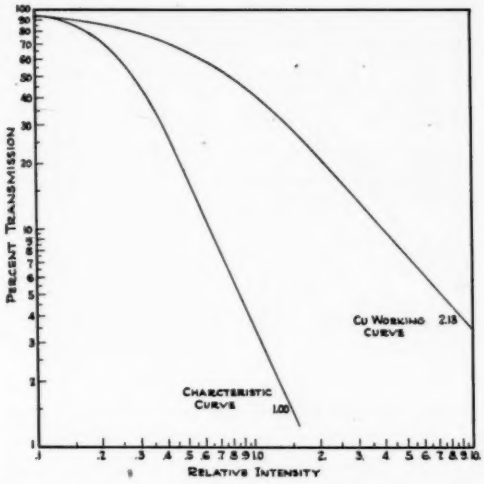

FIG. 3. Typical characteristic curve and copper working curve for steel.

GENERAL PROCEDURE

1. Sample Preparation

Steel samples are drawn by suction from molten metal in 6-mm Pyrex glass tubing, and carefully cooled to prevent chilling; chilled samples cannot be machined. Ends are machined with a special cutting adapter on a small lathe to a 120° angle and two pins are used as electrodes.

Cast iron is poured in a special chill mold forming four 2-in. by $\frac{7}{32}$-in. diameter pins on a large body. The body is used as a sample for chemical analysis. It is important that the sample be solid, free from pinholes, and well chilled. Experience has shown that the best samples are obtained by pre-heating the sample mold before casting the sample. This is easily done by casting an unwanted sample just before the desired sample is poured. Cast iron pins are ground to a 120° angle and used as electrodes as explained above.

2. Exposure

All spectra were photographed in the second order on Eastman Spectrum Analysis No. 1 35-mm film using a Baird[2] three-meter grating

[2] W. S. Baird, "A Practical Grating Spectrograph for Industrial Use," Proc. of the Sixth Summer Conference on Spectroscopy, 1939.

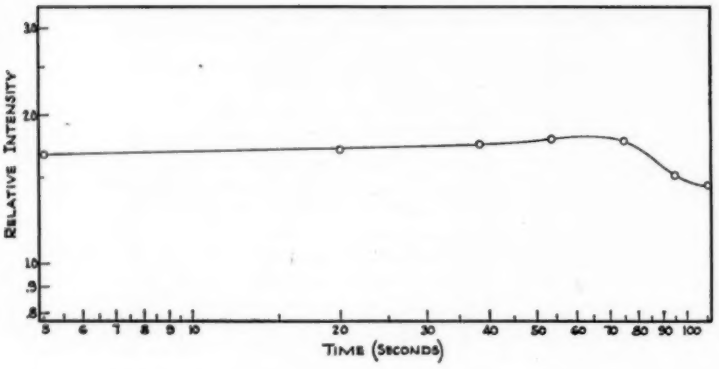

FIG. 4. Burn-out curve—a.c. arc excitation Cu 2824A/Fe 2819A.

spectrograph. Steel and cast iron with less than 0.80 percent copper were given a ten-second pre-spark and exposed photoelectrically about 13 seconds. A rotating sector and filter are used to reduce the light intensity. The sector transmits $\frac{3}{8}$ of the light. The filter is made of $\frac{1}{2}$ mm thickness of Corning 9800 glass and is placed directly over the spectrograph slit. A stream of air at $\frac{1}{4}$ p.s.i. is blown across the sample gap.

Cast iron with greater than 0.80 percent copper is exposed about 16 seconds at 5 kv and $1\frac{1}{2}$ amperes, using an ARL a.c. arc unit.[3] Power is supplied to both spark and arc units by a Burke motor generator set delivering 230 volts regulated by a Westinghouse Silverstat regulator.

The films were processed at a temperature of 68°F as described elsewhere.[4]

3. Data Reduction

The characteristic curve, or $\log T/\log I$ curve, is determined by the two-line method described by Churchill.[5] The characteristic curve of the emulsion is checked frequently, but has been found to change only when a film of a new emulsion number is used.

The films were read using an ARL densitometer[6] which indicates percentage transmission and makes it convenient to plot working curves in terms of log percent transmission.

SPECIFIC PROCEDURE

Steel Containing Greater Than 0.15 and Less Than 0.80 Percent Copper

Both primary and secondary standards[7] were used in working with steel samples. Combining samples from actual heats, National Bureau of Standards and Steel Founders Society of America produced a set of standards which gave good results as to reproducibility and agreement in this range.

Using the ARL spark unit, exposure conditions were $\frac{3}{4}$ kva, $L-1$. This corresponds to a capacity of 0.007 microfarad and an inductance of 0.045 millihenry in the high voltage spark circuit. From each exposure the ratio ICu 3274/IFe 3286 was computed from the characteristic curve. This value was plotted as abcissa and concentration of copper as ordinate on log-log scales. See Fig. 2. From the straight line joining the points, the tangent of the angle α gives the factor by which the horizontal scale of the characteristic curve must be multiplied to obtain the copper working curve for this range. The concentration where the ratio ICu/IFe is unity is the "constant" to be used with this working curve. In actual analysis of an unknown sample, the percent transmission of the iron line is set on the intersection of the vertical scale with the working

[3] H. W. Dietert and J. A. Schuch, "Spectrographic equipment for foundry metal control," Trans. Am. Foundrymen's Assoc. **52**, 889–903 (1945).
[4] W. C. Pierce and N. H. Nachtrieb, Ind. Eng. Chem. Anal. Ed. **13**, 774 (1941).
[5] J. R. Churchill, "Technique of quantitative spectrographic analysis," Ind. Eng. Chem. Anal. Ed. **16**, 653–670 (1944).
[6] H. W. Dietert and J. A. Schuch, "A new projection-comparator densitometer," J. Opt. Soc. Am. **31**, 54–57 (1941).
[7] W. R. Brode and B. F. Scribner, "Report on Standards and Reference Samples for Spectrographic Analysis," Office of Production Research and Development Report No. W-119, p. 3.

curve (in this case the curve with the slope 2.03). The "constant" is indexed on the horizontal scale and the vertical scale is moved to the point at which the transmission of the copper line intersects the working curve: percent concentration of copper is read at this point on the horizontal scale. Typical characteristic curves and copper working curve for this range are shown in Fig. 3. Copper in this range in cast iron is calibrated in the same manner, with the exception that only secondary standards can be used for its calibration as there are no primary standards on the market.

Cast Iron Containing 0.50 to 2.50 Percent Copper

To eliminate the reversal effects of the 3274A copper line caused by spark excitation of high copper concentrations and to attain better reproducibility, the a.c. arc was inspected and found to give best results with 5 kv and 1½ amperes. A burn-out curve of copper with arc excitation (Fig. 4) prompted selection of a thirty-second pre-arc. This gives ample time to stabilize the arc, reach an equilibrium temperature, and average the copper in the sample. The best results also were obtained from the selection of the 2824A copper line with the 2819A iron line as reference. A group of thirty-seven secondary standards with a copper concentration range from 0.18 to 2.59 percent were run repeatedly to check reproducibility. These were plotted as

TABLE II. Comparison of chemical and spectrograph copper analysis in percentages.

Cast iron	Chemical	Spectrograph	Pin size
C-1	0.84	0.80	$\frac{7}{32}$ in.
C-2	1.24	1.24	$\frac{7}{32}$ in.
C-3	1.42	1.41	$\frac{7}{32}$ in.
C-4	1.42	1.54	$\frac{7}{32}$ in.
C-5	2.87	2.60–2.60*	$\frac{7}{32}$ in.
C-6	2.74	3.25–3.26*	$\frac{1}{4}$ in.
Steel			
S-1	0.46	0.44–0.47*	6 mm
S-2	.62	.62– .63*	6 mm
S-3	.63	.63– .65*	6 mm
S-4	.61	.62	6 mm
S-5	.55	.50	6 mm
S-6	.60	.59	6 mm
S-7	.65	.64	6 mm
S-8	.69	.70	6 mm

* Double values are check values required in routine analysis.

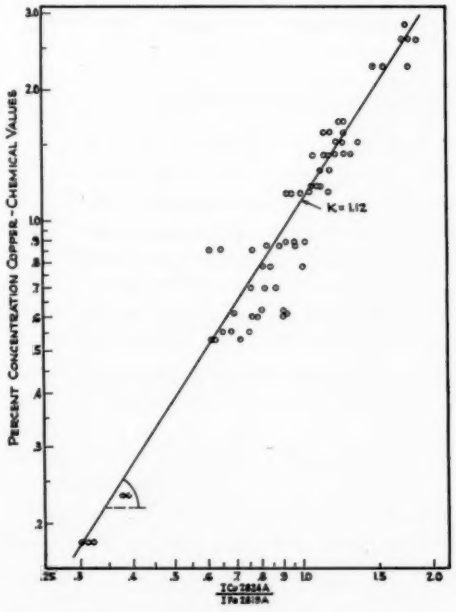

FIG. 5. High copper in cast iron ($\frac{7}{32}$-in. pins) a.c. arc—5 kv.—1.5 amp.

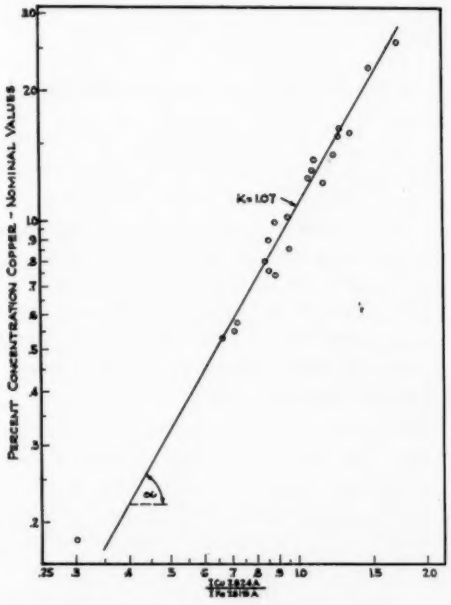

FIG. 6. High copper in cast iron ($\frac{7}{32}$-in. pins) a.c. arc—5 kv—1.5 amp.

described above on log-log paper (see Fig. 5). Those samples which repeatedly checked low or high, in comparison with the majority of the other samples, were given a nominal value to agree with the majority of samples which fell in a straight line and not with previously determined chemical values (Fig. 6). This is in accordance with a precedent set by other authors. Those samples for which values were not in agreement with themselves for several runs were discarded. Cast iron pins both $\frac{7}{32}$-in. and $\frac{1}{4}$-in. diameter were inspected and treated in the same manner. It was found that the $\frac{1}{4}$-in. pins were erratic and plotted on a different slope than the $\frac{7}{32}$-in. pins. The probable reason for this is that the copper is less easily averaged out in the larger pins.

A comparison of analysis determined by chemical methods with that by spectrographic methods is shown in Table II. Results show that the methods described herein give results agreeing within a ± 5 percent limit.

SUMMARY

A method is described in detail for analysis of copper in steel and cast iron with the spark for the range 0.15 to 0.80 percent copper, and in cast iron with the a.c. arc in the range 0.50 to 2.50 percent copper.

ACKNOWLEDGMENTS

Acknowledgment is made to Roy E. Deas for chemical check analysis, to Frank E. Davidson, Jr. for spectrographic assistance, and to Dr. J. H. Coulliette, former spectrographer at the American Cast Iron Pipe Company, who first suggested the use of the a.c. arc for high copper analysis in this laboratory.

A Symposium on Molecular Structure and Spectroscopy

A SYMPOSIUM on Molecular Structure and Spectroscopy will be held at the Mendenhall Laboratory of Physics, at the Ohio State University, from Monday, June 9 through Saturday, June 14. Discussions of the interpretation of molecular spectroscopic data as well as methods for obtaining such data will be held. In addition, sessions will be devoted to microwave and Raman spectroscopy. A dormitory will be available for those who wish to reside on the campus during the meeting. For further information, or for a copy of the program when it becomes available, write to Professor Harold H. Nielsen, Mendenhall Laboratory of Physics, The Ohio State University, Columbus, Ohio.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

The Spectrochemical Analysis of Nickel Alloys*

EDWIN K. JAYCOX
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

A procedure is described for the analysis of nickel alloys for copper, iron, lead, magnesium, manganese, silicon, titanium, and zinc in the range 0.005–0.30 percent and for boron in the range 0.0003–0.03 percent. Samples are taken into solution with dilute nitric acid, evaporated to dryness, and baked at 400°C. The resulting dry nitrate-oxide powder is mixed with pure carbon dust which acts as a buffer and diluent. Aliquots of each sample and of a series of standards are excited in the direct current arc, and their spectra recorded on the same plate. Determinations of the amounts of constituent elements present in the sample are made by measuring the logarithm of the ratio of the relative intensities of a line of the element sought to that of a nickel control line by the general internal control technique.

DISCUSSION

NICKEL and its alloys are used extensively for parts in thermonic devices throughout the communications industry. As in the applications of other materials, the desired properties of nickel alloys depend to a large degree upon their chemical composition, and it is important that this composition be held within strict limits. This requires that control analyses be run on lots of raw material and fabricated parts. Chemical methods are available for the determination of all the usual constituents of nickel alloys, but for most analyses, relatively large amounts of sample are required, which in some cases might be a substantial portion of an entire lot of tiny parts, or in other cases, where only a single small unit is available for test, there would be insufficient sample.

Spectrochemical methods are ideally suited for handling this type of problem with a minimum amount of sample and with moderate precision. The final spectrochemical procedure to be adopted had to be capable of handling samples in the form of fine wires, thin and thick sheets, fabricated parts of all shapes and sizes, and large ingots. Obviously, it was required that all samples be reduced to some common form. The use of solutions or dry powdered salts or oxides fulfill this requirement.

A finely powdered nickel nitrate-oxide mixture, buffered with carbon dust, was found to be the most satisfactory form in which to handle samples and standards for the spectrochemical analysis of nickel alloys for their most common constituents. Nickel solutions, in concentrations of 50 mg/ml or greater, when dried in graphite cup electrodes are difficult to handle. They become very syrupy and have a tendency to climb out of the cups and down the sides of the electrodes. Because of this characteristic the final dried deposits are very non-uniform from electrode to electrode, and poor reproducibility of spectra results upon subsequent excitation. If such solutions are first evaporated to dryness in a Pyrex evaporating dish and baked at 400°C for fifteen minutes, a nitrate-oxide mixture is obtained. This can be readily ground to a fine powder and mixed in an agate mortar until a a very uniform mixture of all constituents is obtained. The addition of 1 to 3 parts of pure carbon dust to one part of the nitrate-oxide mixture aids materially in steadying conditions in the arc, and thereby produces much more uniform spectra than can be obtained without it. The proportions of carbon dust used depend upon the sensitivity and the amount of the element sought. Thus, for the determination of boron at 0.0005 percent and zinc at 0.005 percent, the minimum carbon dust addition is used, because in the first case a very small quantity of boron is sought, and in the second case the element sought, zinc, has relatively low sensitivity. For the determination of the other elements in the ranges covered by the procedure, the larger dilution with carbon dust is used, because of the additional buffering action obtained, as well as the desirability of stretching further very limited size samples.

* Presented at the Thirty-First Annual Meeting of the Optical Society of America, New York, New York, October 4, 1946.

TABLE I.

Element line λ	Control line λ
B 2497.7, 2496.8	Ni II 2505.8, II 2510.9
Cu 3274.0, 3247.5	Ni 2436.7
Mg 2779.8, 2795.5	Ni 2436.7, 2834.6
Mn II 2605.7, II 2593.7, II 2576.1	Ni 2436.7
Fe II 2599.4, 2720.9, 3020.6	Ni 2436.7, 2834.6
Si 2516.1	Ni 2436.7
Ti 3349.4	Ni 2436.7, 2834.6
Pb 2833.1	Ni 2834.6
Zn 3345.0	Ni 2834.6

All quantitative spectrograms are taken through a three step filter, placed in front of the spectrograph slit, which transmits respectively 100, 20, and 4 percent of the light from the source. This permits great latitude in the interpretation of the plates, since the density of any selected line will nearly always lie in a suitable density range in one of the three recorded spectrograms.

APPARATUS

Spectrograph—quartz prism or grating of medium or large dispersion.

Neutral filter or radial sector for exposure control.

Densitometer.

D.c. arc source, 250V-15 ampere capacity.

Plates or film—Eastman process or equivalent.

Aluminum block for holding electrodes.

Fume hood, hot plate, muffle oven.

Dark room facilities.

PROCEDURE

A. Preparation of Standards

All standards except those for silicon and titanium are prepared in the following manner:

To 20-ml aliquots of pure nickel nitrate solution (Ni = 100.0 mg/ml), add measured amounts of the elements sought from standardized nitrate solutions (boron from a boric acid solution) to make a series of standards to cover the desired ranges in steps of factors of 2 or 3. No more than four elements should be included in a single set of standards. The following are suggested compositions and ranges:

Boron: 0.0003–0.03 percent.

Copper, iron, magnesium, and manganese: 0.005–0.3 percent.

Lead and zinc: 0.005–0.30 percent.

Evaporate solutions to dryness. Bake at 400°C for 15 minutes. Grind to a fine powder and mix thoroughly in an agate mortar. Mix with pure carbon dust, one part of carbon dust to one part of standard for boron, lead, and zinc; three parts of carbon dust to one part of standard for copper, iron, magnesium, and manganese.

Silicon and titanium standards are made by adding to 20.0 ml aliquots of pure nickel nitrate solution (Ni = 100 mg/ml), weighed amounts of SiO_2 and TiO_2, calculated as percent of silicon and titanium with respect to nickel. Suggested range 0.005 and 0.10 percent. These oxides will of course not go into solution, but will be incorporated with the dried salts following evaporation and baking as above. To one part of the dry nitrate-oxide, mix three parts of pure carbon dust.

B. Preparation of Samples

Take samples into solution in dilute nitric acid. Approximately 1 g of representative sample should be used if available, but a complete analysis can be run on 200 mg if sample is limited.

Evaporate solution to dryness in a Pyrex evaporating dish. Bake at 400°C for 15 minutes. Grind to a fine powder in an agate mortar. Mix a portion of this powder 1 part to 1 part of pure carbon dust; and a second portion 1 part to 3 parts of pure carbon dust.

C. Electrodes

Lower: Cup, $\frac{1}{8}$″ deep, cut into the end of a $\frac{1}{4}$″ diameter purified graphite rod with a No. 15 drill. Fill each cup level full with its designated sample or standard.

Upper: Flat end, $\frac{3}{16}$″ diameter, purified graphite rod.

D. Excitation and Photography

Determination of boron. Arc the samples, four aliquots, and a series of standards in the d.c. arc at 10.0 amperes for 60 seconds. Record the spectra on an Eastman Process plate or equivalent.

Determination of copper, iron, magnesium, manganese, silicon, and titanium. Arc the samples, four aliquots, and a series of standards in the d.c. arc at 8.0 amperes for 60 seconds. Record the

spectra on an Eastman Process plate or equivalent.

Determination of lead and zinc. Arc the samples, four aliquots, and a series of standards in the d.c. arc at 8.0 amperes for 30 seconds. Record the spectra on an Eastman Process plate or equivalent.

General. Adjust the light entering the spectrograph from the source, by means of a neutral step filter or rotating sector, to bring the density of the spectral lines used in the photometry within a desirable density range (0.2–1.6).

Develop, fix, and dry plate by an accepted photographic processing technique.

E. Photometry

The lines to be used in the photometry are listed in Table I.

Measure the density of the selected lines in both standards and samples.

If the plate has been calibrated, convert densities to relative intensities and from the data obtained on the standards, plot:

Concentration of element sought,

$$X - Vs - \log_{10}(I_x/I_{Ni}),$$

where I is the relative intensity of the lines designated by the subscripts; and determine the amount of element X in the sample from this curve.

If the plate has not been calibrated, plot:

Concentration of element,

$$X - Vs - \log_{10}(D_x/D_{Ni}),$$

where D is the density of the lines designated by the subscripts; and determine the amount of element X in the sample from this curve. This last procedure is permissible if Eastman process plates, or equivalent, are used and if the lines used have densities within the limits of 0.2 and 1.5. Also, since the spectrograms of the standards are on the same plate, the plate is actually calibrated in terms of the standards.

When this procedure is used on a routine basis, the standards need not be run on each plate after the working curves have been firmly established. However, in this case, the calibration of the plate is necessary and all densities must be converted to relative intensities.

TABLE II

Determination of Iron

Sample	No. Determinations	% Fe Spectro	% Fe Chemical	Δ	% Δ
S-64 Cathode	8	0.063	0.063	0.000	0.0
S-65 Cathode	8	0.051	0.05	+0.001	+ 2.0
S-66 Cathode	4	0.067	0.062	+0.005	+ 8.1
J-54 (Soln)	4	0.056	0.060	−0.004	− 6.6
D-70 Cathode	4	0.074	0.07	+0.004	+ 5.7
D-71 Cathode	4	0.076	0.07	+0.006	+ 8.5
A.S.T.M. No. H1400	8	0.036	0.033	+0.003	+ 9.1
				Average	± 5.3

Determination of Copper

Sample	No. Determinations	% Cu Spectro	% Cu Chemical	Δ	% Δ
S-65 Cathode	8	0.027	0.03	0.003	−10.0
S-66 Cathode	4	0.012	0.017	0.005	−29.0*
J-54 (Soln)	4	0.064	0.060	0.004	+ 6.7
D-70 Cathode	4	0.019	0.02	0.001	− 5.0
D-71 Cathode	4	0.018	0.02	0.002	−10.0
A.S.T.M. No. H1400	8	0.011	0.11	0.000	0.0
				Average	± 6.3

Determination of Manganese

Sample	No. Determinations	% Mn Spectro	% Mn Chemical	Δ	% Δ
S-64 Cathode	8	0.100	0.11	0.01	− 9.1
S-65 Cathode	4	0.125	0.124	0.001	+ 0.8
S-66 Cathode	4	0.094	0.11	0.016	−14.5
J-54 (Soln)	4	0.059	0.060	0.001	− 1.7
D-70 Cathode	4	0.110	0.120	0.010	− 8.3
D-71 Cathode	4	0.110	0.115	0.005	− 4.5
A.S.T.M. No. H1400	8	0.128	0.126	0.002	+ 1.6
				Average	± 5.6

Determination of Magnesium

Sample	No. Determinations	% Mg Spectro	% Mg Chemical	Δ	% Δ
J-79-I	4	0.080	0.080	0.000	0.0
J-79-II	4	0.080	0.079	+0.001	+ 1.3
D-70	4	0.096	0.10	−0.004	− 4.0
D-71	4	0.093	0.092	+0.001	+ 1.1
A.S.T.M. No. H1400	8	0.041	0.047	−0.006	−14.0
				Average	± 6.1

Determination of Silicon

Sample	No. Determinations	% Si Spectro	% Si Chemical	Δ	% Δ
D-70	4	0.036	0.033	0.003	9.1
D-71	4	0.031	0.03	0.001	+ 3.3
A.S.T.M. No. H1400	8	0.012	0.013	−0.001	− 7.7
				Average	± 6.7

Determination of Titanium

Sample	No. Determinations	% Ti Spectro	% Ti Chemical	Δ	% Δ
D-70	4	0.028	0.03	−0.002	− 6.7
D-71	4	0.025	0.025	0.000	0.0
A.S.T.M. No. H1400	8	0.033	0.032	+0.001	+ 3.1
				Average	± 3.3

Average % Δ all elements ±5.5.
* Not included in average.

EXPERIMENTAL

The data of Table II indicate the accuracy of the procedure, as compared to routine chemical

methods, on a number of samples. Some of the chemical values were rounded off to one significant figure so the accuracy in some cases might be better than indicated. The average deviation from the chemical values is about ±60 parts/thousand, for copper, iron, magnesium, manganese, silicon, and titanium. No data have been obtained on the accuracy of this procedure for the determination of boron, lead, and zinc, but the precision is of the same order of magnitude as for the other elements.

CONCLUSIONS

This procedure has the advantage of applicability to the analysis of nickel alloys in any form, e.g., wire, sheet, fabricated parts, or ingots. A complete quantitative analysis can be run in triplicate on as little as 200 mg of sample which is a considerable saving in samples as compared to that needed for a complete chemical analysis. Synthetic standards can be readily prepared in the laboratory to include whatever elements and ranges are desired. The speed and precision are moderate.

The author wishes to acknowledge his indebtedness to Mr. G. W. C. Davis and Miss M. M. Hoogstraat for their aid in obtaining some of the spectrochemical data; and to Mr. H. E. Johnson's group who contributed the chemical results.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

Spectrochemical Analysis of Ceramics and Other Non-Metallic Materials*

EDWIN K. JAYCOX
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

The procedure described is applicable to the quantitative spectrochemical analysis of ceramics, ashes, ores, paints, and other non-metallic materials for the determination of most of the common metals and their oxides. These include: aluminum, boron, barium, beryllium, calcium, copper, chromium, iron, lithium, magnesium, manganese, sodium, lead, silicon, titanium, zinc, and zirconium, in the general range of 0.30–70.0 percent. Samples in the form of a fine powder are mixed 1 part of sample to 10–100 parts of a suitable metal oxide which serves as a buffer, diluent, and internal control. Carbon dust is added to this mixture for its additional buffering effect. Spectra are obtained of the samples and of an appropriate series of standards. Determinations of the amount of element sought are made, in most cases by the well known internal standard technique, in others by the simple comparison standard procedure.

DISCUSSION

THE quantitative analysis of ceramics, ores, slags, ashes, paints, and other non-metallic materials, for major and minor metallic ion components, is often a long and tedious job when done by classical chemical procedures. Relatively large samples are required and in many cases insufficient sample is available to run a complete analysis. Interference of one element with another is often a source of trouble.

The spectrograph is not generally thought of as being a suitable tool for the determination of major constituents in materials, but has nevertheless been successfully applied in this range in our laboratories, for a number of years, particularly for the analysis of copper base alloys,[1] ceramics, solders, Alnico, and ashes. Oshry, Ballard, and Schrenk[2] have determined silicon, iron, and aluminum at high concentrations in mineral powders. Fitze and Murray[3] have also employed it with good results for the analysis of small samples of miscellaneous materials.

When samples are excited directly, the densi-

* Presented at the Seventh Pittsburgh Conference on Applied Spectroscopy, Mellon Institute, Pittsburgh, Pennsylvania, November 15, 1946.

[1] E. K. Jaycox, J. Opt. Soc. Am. **35**, 175–179 (1945).
[2] H. I. Oshry, J. W. Ballard, and H. H. Schrenk, J. Opt. Soc. Am. **32**, 672–680 (1942).
[3] E. J. Fitze and W. M. Murray, Ind. and Eng. Chem., Anal. Ed. **17**, 145–147 (1945).

ties of the principal lines of the major components are usually too high for accurate densitometry. Furthermore, at high concentrations, self-absorption becomes an important factor, causing many of the lines of most elements to become entirely unsuitable for quantitative work. The use of filters or other devices for diminishing the amount of light admitted to the spectrograph does not help in these cases, since the phenomena of self-absorption takes place in the arc column itself. The reduction of the concentration of the element in question, by the addition of a buffer material, eliminates most of the troubles arising due to self-absorption, and at the same time reduces the densities to measurable ranges. Buffering also reduces the complexity of the spectrograms by eliminating many of the less sensitive lines and, in addition, reduces the interinfluence of component elements upon each other in complex and variable matrices.

This general principal of employing a buffer diluent may be applied to samples in the form of powders, solutions or low melting point alloys. In the latter case a pure metal may be used as the diluent and analyses run using self-electrodes cast from the melt.

Ceramics and other insoluable materials are most conveniently handled as powders. A representative sample is ground to a fine powder and mixed with a suitable buffer material. Lead monoxide, bismuth trioxide, cupric oxide, nickel oxide,[2] and barium nitrate[3] have been used successfully as a buffer and internal control. Of these, the black cupric oxide (CuO) has been found, by us, to be very satisfactory. The basic mixture usually employed is one part of sample to 19 parts of pure cupric oxide, to which is added for additional buffering, 40 parts of pure carbon dust. Such a mixture is satisfactory for the determination of most elements in the range of 0.30–35.0 percent. When it is desired to determine an element in any amount greater than 35 percent, it is better to use a mixture in the ratio of 1:39:80. A ratio of 1:9:20 is sometimes used when the concentration of the element sought is less than one percent. Standards are always made up in the ratio of 1:19:40, and a factor used in the final calculations when samples have been prepared in a different ratio.

Since such relatively large quantities of buffers are used, it is of primary importance that they be very pure, because in the final mixture, an impurity of 0.01 percent would be equivalent to 20 or 40 times that amount, or .2–.4 percent. For this reason all buffer materials are carefully checked, and are not used if they contain greater than 0.001 percent of the element sought in the sample.

TABLE I.

Element line—A	Control line—A
Al—3082.2, 3092.7	Cu—2768.9
B—2497.7, 2496.8	Cu—2768.9
Be—3221.3, II 3130.4, II 3131.1	Cu—2768.9
Ba—II 4934.1, 4554.0	Cu—4249.0
Ba—3071.6	Cu—3126.1
Ca—4226.7	Cu—4249.0
Ca II—3158.9	Cu—3126.1
Cr—4254.3, 4274.8, 4289.7	Cu—4249.0
Fe—II 2599.4	Cu—2768.9
Li—3232.6	Cu—3126.1
Li—4602.9	Cu—4249.0
Mg—2798.1, 2783.0, 2776.7	Cu—2768.0
Mn—2798.3, II 2576.1	Cu—2768.9
Na—5895.9, 5890.0	No control
Na—3302.3, 3303.0	Cu—2768.9
Pb—2873.3, 2833.1	Cu—2768.9
Si—2506.9, 2514.3, 2516.1, 2524.1	Cu—2768.9
Ti—II 3349.0, 3361.2	Cu—3126.1
Zn—3345.6	Cu—3126.1, 2768.9
Zn—4810.5	Cu—4249.0
Zr—II 2571.4	Cu—2768.9

APPARATUS

Spectrograph—quartz prism or grating of medium or large dispersion.
Neutral filter or radial sector for exposure control.
D.c. arc source—250 v—15 ampere capacity.
Plates—Eastman process, process panchromatic or equivalent.
Densitometer.
Agate mortar.
Aluminum block for holding electrodes.
Dark room facilities.

PROCEDURE

A. Preparation of Standards

National Bureau of Standards certified samples, or other samples of known composition, may be used as a base for preparing spectrochemical standards. Standards may also be prepared by mixing accurately weighed amounts of the desired pure compounds. Bake all ingredi-

TABLE II.

Sample	% Compound	% Spectro	Chemical	Δ	% Δ
F-66	Al_2O_3	7.61	7.58	+0.03	+0.4
S-7	Al_2O_3	1.04	1.0	+0.04	+4.0
Sierramic Talc	Al_2O_3	1.07	1.15	−0.08	−7.4
N.B.S. No. 76	Al_2O_3	36.2	37.7[1]	−1.5	−4.0
F-66	BaO	16.4	15.74	+0.76	+4.8
S-7	BaO	4.13	4.17	−0.04	−1.0
F-66	CaO	0.94	0.54	+0.30	+47.0*
S-7	CaO	0.64	none	—	—*
Sierramic Talc	CaO	0.42	0.43	−0.01	−2.3
N.B.S. 76	CaO	0.265	0.27[1]	−0.005	−1.8
F-66	Fe_2O_3	1.15	0.91	+0.24	+26.0*
S-7	Fe_2O_3	0.71	0.66	+0.05	+7.5
Sierramic Talc	Fe_2O_3	1.37	1.39	−0.02	−1.4
N.B.S. No. 76	Fe_2O_3	2.30	2.37[1]	−0.07	−3.0
F-66	MgO	23.7	24.6	−0.9	−3.7
S-7	MgO	50.2	50.9	−0.7	−1.4
Sierramic Talc	MgO	31.6	31.2	+0.4	+1.3
F-66	SiO_2	47.1	50.12	−3.02	−6.0
S-7	SiO_2	44.5	42.99	+1.51	+3.5
Sierramic Talc	SiO_2	57.2	58.8	−1.6	−2.7
N.B.S. No. 76	SiO_2	56.0	54.7[1]	+1.3	+2.4
Ash-Tar	ZnO	0.53	0.55	−0.02	−3.6
				Av. %Δ	±3.3

* Values omitted from average.
[1] N.B.S. certified values.

ents at 120°C for at least one hour to assure that all absorbed water has been driven off.

Mix thoroughly in an agate mortar, 1 part of standard base with 19 parts of black cupric oxide (CuO), and 40 parts of pure carbon dust, all by weight. Store in screw cap vials.

Single sets of standards may be made to contain several elements, but it is recommended that not more than four elements or compounds be included. Values of standards should be chosen to straddle the anticipated amount of the element expected in the sample as determined by a qualitative analysis.

The elements to be included and their ranges will of course depend upon the problem at hand.

B. Preparation of Samples

Pulverize samples to a fine powder in a suitable mortar. (Agate preferred.) Bake at 120°C in an oven for one hour to expell any absorbed water.

Mix thoroughly, 1 part of sample (100 mg of representative sample) with 19 parts of cupric oxide (CuO) and 40 parts of carbon dust. This is the standard mixture. Other ratios may be used when desired.

Keep samples in tightly capped vials.

TABLE III. Spectrochemical analysis of unknown ceramics.

	1	2	3	4	5	6
BaO	8.3%	5.6%	6.8%	4.0%	6.7%	—
CaO	0.6	0.6	0.6	<0.2	<0.2	—
MgO	31.0	25.0	29.4	24.8	25.4	—
Al_2O_3	3.3	3.8	5.6	4.1	8.0	29.6
Fe_2O_3	1.1	1.4	1.6	1.3	1.7	0.34
SiO_2	54.6	59.0	56.1	64.0	51.0	65.7
TiO_2	<1.0	<1.0	<1.0	<1.0	<1.0	<1.0
Total	99.9%	96.2%	101.1%	96.5%	94.0%	96.6%
	7	8	9	10	11	12
BaO	<1.0%	<1.0%	<1.0%	—	7.6%	5.8%
CaO	1.0	1.0	<1.0	0.2%	0.4	0.3
MgO	25.9	29.6	<1.0	1.0	30.0	29.0
Al_2O_3	7.4	3.2	1.3	1.7	3.8	3.2
Fe_2O_3	1.6	1.0	<1.0	0.13	1.7	1.5
SiO_2	54.0	64.4	16.8	4.0	51.0	53.0
TiO_2	<1.0	<1.0	75.0	80.0	5.3	5.0
Total	91.9%	101.2%	97.1%	87.0%	99.8%	97.8%

C. Electrodes

Lower: Cup, $\frac{1}{8}''$ deep, cut with a No. 15 drill, into the end of a $\frac{1}{4}''$ diameter purified graphite rod. Fill each cup level full with its designated standard or sample.

Upper: Flat end, $\frac{3}{16}''$ diameter, purified graphite rod.

D. Excitation and Photography

Arc the samples and a series of standards in the 250 v d.c. arc at 10 amperes for 60 seconds. Record the spectra on an Eastman process or process panchromatic plate, or equivalent. The process emulsion is preferred in all cases where it is not necessary to use lines having wave-length greater than 4000A. For higher wave-lengths the process panchromatic must be used, and may be used for lines at lower wave-lengths.

Adjust the light entering the spectrograph from the source, by means of a neutral step filter, rotating sector or variable aperture, to bring the densities of the lines to be used in the photometry, within a desirable density range.

Develop, fix, and dry plate by an accepted photographic processing technique.

E. Photometry

In Table I are listed the lines used for the determination of the amounts of each constituent

element present. In the first column are shown the line or lines that can be used for the element sought. In the second column, the copper control line is shown, where a control is used.

Measure the density of the selected lines in the spectra of both standards and samples. If the plate has been calibrated, convert densities to relative intensities, and from the data obtained on the standards, plot a curve of:

Concentration of Element X-*vs.*-$\log_{10}(I_x/I_{cu})$,

and determine the amount of x in the sample from this curve.

If the plate has not been calibrated, plot:

Concentration of Element X-*vs.*-$\log_{10}(D_x/D_{cu})$,

where D is the density of the lines designated by the subscripts, and determine the amount of x in the sample from this curve. This procedure is permissible where spectrograms of the standards are taken on the same plate with those of the samples, and if the lines used have densities within the limits of 0.2 and 2.0.

EXPERIMENTAL

The data in Table II indicate the accuracy to be expected of this procedure. Nineteen out of twenty-two comparisons with chemical determinations are in good agreement. Three are not. These have been omitted from the average. These data do not cover all of the elements and compounds embraced by this procedure but our experience with miscellaneous applications indicates that the precision is, in general, better than ± 100 parts/thousand.

In Table III is shown the complete spectrochemical quantitative analysis of twelve ceramic samples whose composition was entirely unknown to us. Qualitative analyses were first run, followed by quantitative analyses for all elements estimated to be greater than 0.5 percent. The total sum of the amounts of compounds found should add up in each case to 100 percent $\pm$ the precision of the procedure. The average deviation from 100 percent for these twelve samples is about 4 percent which agrees quite well with the data of Table II.

CONCLUSIONS AND ACKNOWLEDGMENTS

This procedure has wide applicability. Any inorganic material that can be pulverized to a fine powder is subject to analysis. Complete analyses can be run on as little as 10 mg of sample. Ashes from organic samples may also be analyzed using the fundamental precepts of this method; however, the details of techniques for preparing samples differ considerably and for this reason their analyses have not been included in this presentation.

The author wishes to acknowledge his indebtedness to Mr. G. W. C. Davis and Miss M. M. Hoogstraat for their work in obtaining some of the spectrochemical data, and to Mr. H. E. Johnson's group for the chemical analyses.

The Relation of the Configuration of a Brightness Contrast Border to Its Visibility

GLENN A. FRY*
School of Optometry, The Ohio State University, Columbus, Ohio
(Received October 14, 1946)

PART I

Introduction

A BRIGHTNESS contrast border occurs wherever two areas adjoining each other in the visual field have no other means of demarcation than the brightness difference between them. If the brightness difference (ΔB) between the areas on the two sides of the border is gradually reduced, it is possible to find a minimum difference at which the border is visible and this is called the brightness difference threshold for a brightness contrast border.

An attempt has been made in the investigations described in this paper to isolate several variables in the external stimulus pattern and distribution of intensity on the retina, which appear important for the visibility of a border. In general four types of patterns have been investigated.

1. Long straight borders.
2. Circular borders.
3. Segmented borders composed of straight line segments.
4. Wavy borders composed of alternating convex and concave arcs.

Attention has been paid to the following variables.

1. The length of a straight segment in a border.
2. The length of an adjoining straight segment.
3. The angle between adjoining straight segments.
4. The curvature of a border.
5. The area inclosed by a border.

Some attention has also been paid to the fact that near the threshold level a border having a complex configuration will be transformed to a simpler configuration. For example, a saw-tooth or wavy border may be transformed to a straight line, and a regular polygon may be transformed to a circle.

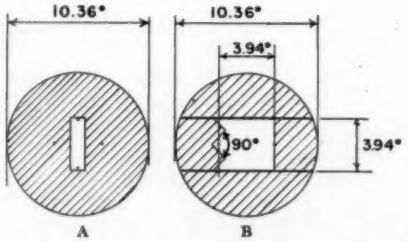

FIG. 1. Stimulus patterns for investigating the visibility of rectangles, segmented, and long straight borders.

* Aided by a Grant for Research in Physiological Optics by the Bausch and Lomb Optical Company.

PART II. THE RELATION OF THE LENGTH OF A BORDER TO ITS VISIBILITY

Stimulus Patterns

This part of the paper is concerned with two types of stimulus patterns for measuring the contrast threshold. These are illustrated in Fig. 1. The one in Fig. 1A is a rectangle on a slightly darker background.

The other (Fig. 1B) is more complicated; the two black lines across the background eliminate the possibility of treating the central bright area as a simple single test object with a single threshold, and make it possible to measure separately the threshold for the saw-tooth border on the left and the straight border on the right. The segments in the saw-tooth contour join at right angles. The length of the segments can be changed by using a larger or smaller number of teeth. In the experiment described below, the following combinations of number of teeth and length of segments were used:

Number of teeth	2	4	8	16
Length of segment (visual angle subtended at first nodal point of the eye)	83′	41.5′	20.7′	10.4′

Apparatus

The apparatus for the rectangular type of test object is illustrated in Fig. 2. The circular aperture (A) subtends 10.36° at the first nodal point of the eye. It is covered by a sheet of opal-flashed

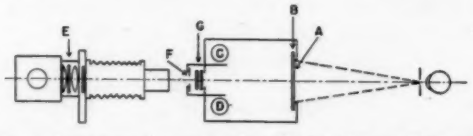

FIG. 2. Apparatus for measuring the brightness difference threshold of a brightness contrast border.

TABLE I. Contrast thresholds ($\Delta B/B$) for rectangles of various widths.

Subject	Width of rectangle (minutes)	Long sides vertical		Long sides horizontal		Average
		Descending	Ascending	Descending	Ascending	
G.A.F.	10.4	.0575	.0488	.0485	.0440	.0497
	20.7	.0564	.0443	.0490	.0488	.0496
	41.5	.0517	.0526	.0415	.0534	.0498
	83.0	.0530	.0455	.0575	.0520	.0520
	236.0	.0495	.0561	.0510	.0530	.0524
J.O.F.	10.4	.0380	.0317	.0324	.0271	.0323
	20.7	.0356	.0378	.0324	.0296	.0339
	41.5	.0306	.0358	.0355	.0342	.0340
	83.0	.0315	.0340	.0323	.0324	.0326
	236.0	.0440	.0408	.0456	.0396	.0423

glass which is illuminated by two lamps, C and D. The brightness, B, was 77 c/m². The subject fixated an imaginary point at the center of a constellation of four small ink points on the opal-flashed glass (Fig. 6A).

A lantern slide projector was used for projecting on the opal-flashed glass an image which served as the test object. The center of the image coincided with the point of fixation. The slides consisted of black cardboard with apertures cut to the shape and size required for the test object.

A plate of ground glass (E) was inserted between the two condenser lenses of the projector so that an iris diaphragm, F, could be used to vary the brightness (ΔB) of the projected image. A pair of crossed Polaroids, G, was used to vary ΔB for making the threshold measurements, and the iris diaphragm was adjusted at the beginning of the experiment to make the Polaroid settings fall in the middle region of the scale.

An artificial pupil 3.08 mm in diameter was mounted at a distance of 780 mm from the milk glass, and the nodal point at the eye was approximately 20 mm behind this.

The subject controlled the Polaroid setting by a knurled knob. Starting at some randomly selected sub-threshold value, he gradually increased the brightness of the image till it became visible. After each setting, the subject removed his hand from the knob and the operator reset the Polaroids to a sub-threshold value.

The same apparatus was used for the saw-tooth test object with the exception that a second piece of milk glass was used, on which two black lines one millimeter thick were painted and also two fixation points, one at the center of the straight side of the test object and the other at the center of the saw-tooth border.

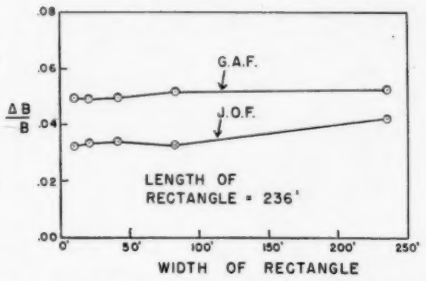

FIG. 3. The effect of varying the width of a rectangle upon its visibility.

Effect of Varying the Width of a Rectangle

The effect of varying the width of a rectangular test object was investigated by using rectangles which have the same length (236′) but different widths as follows: 10.4′, 20.7′, 41.5′, 83′, and 236′. In the first series of measurements, the long sides of the rectangle were oriented in a vertical direction. The different widths were tried one at a time starting with the widest and going down the series (descending order). Five settings were made for each size.

A second series of measurements was made starting with the narrowest rectangle and going up the series (ascending order). A third and a fourth (ascending and descending) series of measurements were made with the long sides oriented horizontally.

The results for two subjects are shown in Table I. The four series of measurements showed no significant differences, and the results were averaged as shown in the fifth column of Table I and are plotted in Fig. 3.

Subject G.A.F. used his right eye with an occluder in front of the left eye. A cylinder $-.50\times120$ was mounted between the artificial pupil and the eye for compensating ametropia. Subject J.O.F. also used his right eye and a combination of a sphere and a cylinder ($+.25$ $-.50\times100$) for compensating his ametropia. One can conclude from these data that varying the width from 10.4′ to 236′ does not effect the threshold appreciably.

In previous papers,[1,2] it has been shown that

[1] G. A. Fry and P. W. Cobb, "New method of determining the blurredness of the retinal image," Trans. Am. Acad. Ophth. and Oto., 432–448 (1935).

[2] G. A. Fry, The Optometric Weekly **37**, 1521–1523, 1537 (1946).

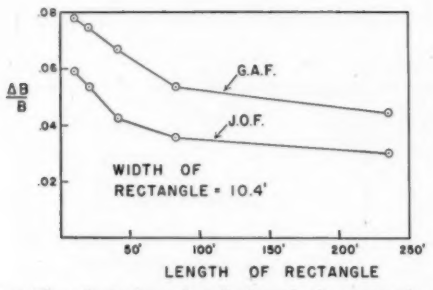

FIG. 4. The effect of varying the length of a rectangle upon its visibility.

when the width is reduced to a size commensurate with the blur circle for a point source, further decrease in the width does affect the threshold, but this effect can be attributed to the fact that the blurredness of the retinal image causes a reduction in contrast.

One may draw the general conclusion, that so long as the width exceeds the diameter of the blur circle for a point source, the threshold is independent of the width. As pointed out in the previous paper, this conclusion is intended to apply only to foveal stimuli. In the periphery of the retina, the convergence of a number of photoreceptors on a single ganglion cell makes the problem of visibility in the periphery one of a different order.

One can conclude also from this experiment that sheer increase in area (i.e., number of photoreceptors exposed to the test object) will not affect the threshold.

As the width of the rectangle increases, the sides fall on more and more peripheral portions of the retina, but this variable does not seem to affect the results. This factor would be expected to produce a marked effect if the brightness of the background were reduced to the threshold level for cone vision. Hence in order to prevent the distribution of rods and cones from influencing the results, the highest brightness level attainable with the apparatus was used.

One can also conclude from the experiment that varying the length of the short sides of a rectangle does not affect its visibility.

Effect of Varying the Length of a Rectangle

The series of rectangles used in this experiment had a constant width of 10.4′ and different lengths were used as follows: 10.4′, 20.7′, 41.5′, 83′, 236′.

The same procedure was used as in the preceding experiment. The results which are shown in Table II and Fig. 4 indicate that the length of a rectangle does affect the threshold. Since it has already been shown that increase in area does not affect the threshold, one must conclude that the determining factor is the length of the two long borders.

Since the short sides become visible at the same time as the long sides, one must conclude that the visibility of the short sides is affected by the length of the long sides.

Effect of Varying the Length of the Segments in a Saw-Tooth Contour

The saw-tooth pattern provides a more satisfactory method of demonstrating the dependence of the visibility of a border upon its length, because area as a variable is eliminated.

The experiment was carried out by starting with the finest saw-tooth pattern. With this particular pattern, two thresholds were obtained for the saw-tooth contour, one at which the whole contour was perceived as a straight line, and a second threshold at which the individual saw teeth were visible. The operator set the value of ΔB at a sub-threshold level, and the subject fixated the fixation point on the left, and gradually increased ΔB till the "straight border" was visible and then further increased it till the individual teeth were visible. The subject then removed his hand from the control knob, and the operator reset the value of ΔB to a sub-threshold level chosen at random. Five such sets of readings were made. The subject then fixated the point on the right, and five settings were made for the threshold for the straight edge going from invisibility to visibility in each case.

TABLE II. Contrast thresholds ($\Delta B/B$) for rectangles of various lengths.

Subject	Length of rectangle (minutes)	Long sides vertical		Long sides horizontal		Average
		Descending	Ascending	Descending	Ascending	
G.A.F.	10.4	.0805	.0883	.0715	.0722	.0781
	20.7	.0786	.0815	.0645	.0750	.0749
	41.5	.0670	.0669	.0711	.0588	.0670
	83.0	.0593	.0551	.0570	.0449	.0540
	236.0	.0570	.0502	.0496	.0430	.0450
J.O.F.	10.4	.0575	.0540	.0656	.0625	.0599
	20.7	.0572	.0545	.0500	.0534	.0538
	41.5	.0488	.0495	.0366	.0362	.0428
	83.0	.0422	.0420	.0248	.0332	.0356
	236.0	.0312	.0333	.0253	.0324	.0306

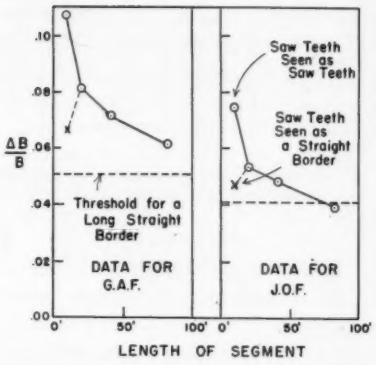

FIG. 5. The effect of varying the length of the segments in a saw-tooth border upon its visibility.

The next coarser saw-tooth pattern was then substituted for the original pattern. With this pattern and for all the coarser saw-tooth patterns only one threshold was obtained, i.e., individual saw teeth were plainly visible whenever the contour was visible at all. With this pattern, five threshold measurements were made for the saw-tooth edge and five for the straight edge. The same procedure was followed for the next coarser saw-tooth pattern and for the coarsest and then the whole experiment was repeated backward.

The results for the two halves of the experiment (Table III) showed no significant differences and were averaged. They are plotted in Fig. 5. In the case of the finest saw-tooth pattern, the cross represents the threshold value at which the saw-tooth contour was first visible as a

TABLE III. Contrast thresholds ($\Delta B/B$) for saw-tooth segments of different length.

Subject	Length of segments (minutes)	Ascending Saw-tooth	Ascending Straight edge	Descending Saw-tooth	Descending Straight edge	Average Saw-tooth	Average Straight edge
G.A.F.	10.4 (1st)	.0610	.0418	.0716	.0610	.0663	.0514
	10.4 (2nd)	.1030		.1131		.1080	
	20.7	.0800	.0480	.0838	.0562	.0819	.0521
	41.5	.0635	.0434	.0805	.0505	.0720	.0469
	83.0	.0650	.0541	.0580	.0521	.0615	.0531
Average							.0509
J.O.F.	10.4 (1st)	.0563	.0394	.0380	.0449	.0472	.0422
	10.4 (2nd)	.0860		.0635		.0748	
	20.7	.0567	.0421	.0502	.0338	.0535	.0380
	41.5	.0458	.0353	.0508	.0390	.0483	.0372
	83.0	.0390	.0412	.0396	.0470	.0393	.0441
Average							.0404

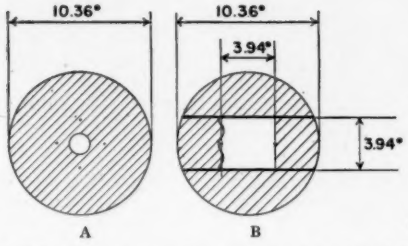

FIG. 6. Stimulus patterns for investigating the visibility of curved, wavy, and long straight borders.

straight line, and the open circle represents the average threshold value at which the individual saw teeth are visible.

The results show that the threshold for the straight edge was approximately constant during the experiment, and the differences in the threshold values for segments of different length show that the length of the segment definitely affects the threshold. It is pointless to try to carry the experiment to segments much shorter than the shortest one used in the experiment because as soon as the length of the segment approaches the dimensions of the blur circle of a point source the experiment loses its significance.

The fact that at the threshold level, a saw-tooth pattern will be transformed by the visual mechanism to a straight edge, when the length of the segments is a number of times larger than the diameter of the blur circle, constitutes a striking demonstration that the mechanism for receiving and transmitting the impression of a border is much more adapted for long straight borders than for short ones.

PART III. THE RELATION OF THE CURVATURE OF A BORDER TO ITS VISIBILITY

Stimulus Patterns

This part of the paper is concerned with two types of stimulus patterns for measuring the contrast threshold. These are illustrated in Fig. 6. The one in Fig. 6A is a disk on a slightly darker background. The other (Fig. 6B) is similar to Fig. 1B except that a wavy border is substituted for the saw-tooth border. The wavy line is a series of convex and concave arcs each of which represents a quarter of a circle. A pair of these arcs constitutes a cycle. In the experiment described below, the curvature of the arcs was

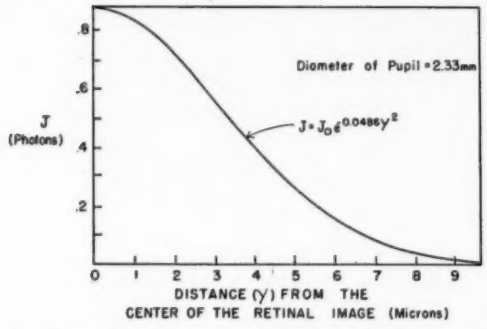

FIG. 7. Distribution of intensity in the retinal image of a small rectangular element of area ($d\rho^2$).

varied by using a larger or smaller number of cycles. The following combinations of radius of curvature and number of cycles were used:

Number of Cycles	2	4	8	16
Radius of Curvature	41.5′	20.7′	10.4′	5.2′

Relation to Previous Experiments

A number of previous investigators have used a disk-shaped test object for investigating the effect of the size of a test object upon the brightness difference threshold. A number of references to the literature can be found in Wald's report.[3] The problem is somewhat different when the test object is fixated centrally as compared with fixating a point to one side so that the disk-shaped test object falls on the periphery of the retina. The problem in the peripheral retina is complicated by the fact that rods are present, and the fact that the distribution of ganglion cells is less concentrated than the distribution of the rods and cones and therefore the pathways for a number of rods and cones must converge upon a single ganglion cell. The fovea is rod-free and also throughout the fovea the population of ganglion cells is considered to be equal to that of the cones. In the present paper, attention is confined entirely to test objects which are fixated centrally. All the investigators have found that the threshold decreases as the size of a centrally fixated disk increases.

Wald[3] has derived an equation for the relationship between the contrast threshold and the area of a disk-shaped test object, which is based upon

[3] G. Wald, "Area and the visual threshold," J. Gen. Physiol. 21, 269–287 (1937–38).

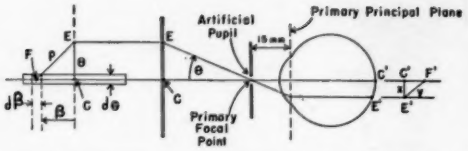

FIG. 8. Geometry involved in calculating the distribution of intensity across the retinal image of a bar of infinitesimal width ($d\theta$).

the assumption that the threshold depends upon the amount of area included within the image of the test object.

In order to avoid confusion, the problem should be divided into two parts involving two ranges for the size of the test object. In the lower range, the diameter of the test object is of the same order of magnitude as the diameter of the blur circle of a point source of light. In this range, changing the size of the test object affects the intensity at the center of the retinal image. As the size of the test object approaches that of a point, the retinal image becomes equivalent so a blur circle for a point source and changing the size produces the same result on the retina as changing intensity. As long as the diameter of the disk is many times the size of the diameter of a blur circle of a point source, changing the diameter of the test object produces a corresponding change in the diameter of the retinal image, but does not affect the intensity at the center of the image.

An equation can be derived for the relation between the intensity at the center of the retinal image and the size of the disk. Let us assume that the curve in Fig. 7 represents the distribution of intensity in the retinal image of a small rectangular element of area which subtends a solid angle $(d\rho)^2$ at the entrance pupil of the eye. It is assumed that this distribution is radially symmetrical. For convenience, J is defined as intensity of retinal illumination (photons) per unit of brightness (c/m²) per unit of solid angle (square minutes). The symbol y represents the distance from the center of the distribution on the retina. In spite of difficulty in defining intensity of retinal illumination and retinal stimulation, the photon as defined by Troland is a perfectly satisfactory unit for indicating relative intensities when a small constant size artificial pupil is used.

The end result is important enough to interrupt the main argument long enough to show that the distribution for a point source in Fig. 7 represents a degree of blurredness which is equivalent to that for a linear source, which Fry and Cobb[1,2] have shown to exist for a narrow vertical line as seen by Fry's right eye through a 2.33-mm artificial pupil.

In Fig. 8, C represents a point at the center of a narrow bright bar which subtends an angle $d\theta$ at the primary focal point of the eye. C' represents the point in the retinal image corresponding to C. E' represents a point at a distance x from the center of the retinal image of the bar. The bar can be subdivided into a number of elements $d\beta$ long and $d\theta$ wide. F represents the center of one of these elements and F' represents the center of the retinal distribution produced by the element. The intensity of illumination (BJ) at E', produced by the element of area, can be obtained from Fig. 7 since

$$\theta = 0.23x, \tag{1}$$

and

$$y = 4.35\rho = (\beta^2 + \theta^2)^{\frac{1}{2}}, \tag{2}$$

when x and y are expressed in micra and ρ, β, and θ are expressed in minutes.

The total intensity (BH) at E' is equal to the sum of all the BJ values for the various subdivisions of the bar. The equation for (BH) is as follows

$$BH = 2B\int_0^\infty J d\beta. \tag{3}$$

The computed values of H for various values of x are shown in Fig. 9. This is identical with the distribution for a narrow bar determined experimentally.

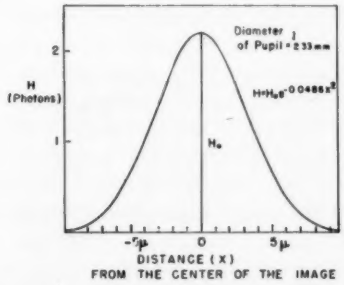

FIG. 9. Distribution of intensity across the retinal image of a bar of infinitesimal width ($d\theta$).

In Fig. 10, C represents a point at the center of a disk of finite size, and C' represents the geometrical image of C on the retina. The disk can be divided into a number of zones of infinitesimal width ($d\theta$), and each zone can be subdivided into a number of elements of area $(d\rho)^2$. The intensity which one of these zones produces at C', is $BJ2\rho d\rho$ in which J is a function of ρ. The value of J for a given value of ρ can be obtained by finding the value of y_{C_1} corresponding to ρ, and then determining the value of J from Fig. 7. The value of $y_{C'}$ corresponding to a given value of ρ can be computed since,

$$\rho = 0.23y_{C'}. \tag{4}$$

The total intensity ($I_{C'}$) at C' contributed by all zones is given by the following formula in which ρ' represents the radius of the disk.

$$I_{C'} = B\int_0^{\rho'} J\rho d\rho. \tag{5}$$

For small values of ρ', J may be assumed to be constant and equal to J_0 and hence

$$I_{C'} = BJ_0\pi(\rho')^2. \tag{6}$$

If for disks of different sizes, the value of $I_{C'}$ for the brightness difference threshold must be a constant,

$$B_{\text{threshold}} = k'/(\pi\rho^2). \tag{7}$$

This means that for small disks the brightness difference threshold must be inversely proportional to the area of the disk, and this is found by experiment to be the case.[4]

The solid curve in Fig. 11 which is a plot of Eq. (5) shows that when the radius of the disk exceeds the radius of the blur circle of a point source, increase in the size of the disk no longer affects the intensity at the center of the image. If

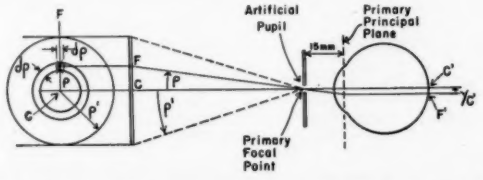

FIG. 10. Geometry involved in computing the intensity at the center of the retinal image of a disk.

[4] G. M. Byram, "Physical and photochemical basis of visual resolving power," J. Opt. Soc. Am. **34**, 571–591 (1944).

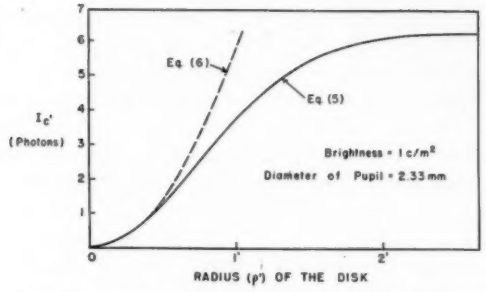

FIG. 11. The effect of varying the size of a disk upon the intensity at the center of the retinal image. Equation (5) is exact, but Eq. (6) is approximate and holds only for disks of small radius.

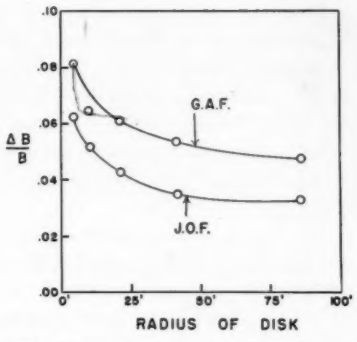

FIG. 12. The effect of varying the size of a disk upon its visibility.

contrast between the intensity at the center of the retinal image and the intensity in the region surrounding the image were the only factor affecting the threshold, one might expect that the value for the threshold would gradually become less and less affected by increase in area, and finally become independent of change in area when the radius of the disk exceeds the radius of the blur circle for a point source. However, experiment shows that the threshold varies with area for much larger disks than those which are of the same order of magnitude as the blur circle of a point source. This means that some other factor must be taken into consideration.

Results which have been obtained with rectangular test objects which are described in the first part of this paper, prove rather conclusively that sheer increase in area included within the retinal image of the test object does not affect the visibility of the border. One factor which might affect the visibility of the border of a disk-shaped test object which does not appear to have been taken into consideration is the curvature of the border. The purpose of the present investigation is to demonstrate the role played by this factor.

TABLE IV. Contrast thresholds ($\Delta B/B$) for disks of different size.

Subject	Order of test objects	Radius of curvature 5.2′	10.4′	20.7′	41.5′	83.0′
G.A.F.	Descending	.0725	.0590	.0540	.0502	.0397
	Ascending	.0914	.700	.0680	.0561	.0553
	Average	.0820	.0645	.0610	.0532	.0475
J.O.F.	Descending	.0625	.0442	.0376	.0346	.0317
	Ascending	.0620	.0586	.0475	.0353	.0342
	Average	.0623	.0514	.0426	.0350	.0330

Apparatus

The apparatus used in this investigation is the same as that described in Part I, except that different slides were used to produce disk type test objects and wavy borders illustrated in Fig. 6.

The Effect of Varying the Diameter of a Disk-Shaped Test Object

The series of disks used in this experiment had diameters which subtended visual angles at the first nodal point of the eye as follows: 10.4′, 20.7′, 41.5′, 83′, and 166′.

The different disks were tried one at a time starting with the largest and going down the series (descending order). Five settings were made for each size, and then the experiment was repeated backwards (ascending order). The results are given in Table IV, and the averages are plotted in Fig. 12. As found by previous investigators, the size of the disk affects its visibility. The smallest disk used subtends a visual angle which is much larger than the diameter of a blur circle and therefore the results cannot be attributed to the effect of blurredness upon the intensity at the center of the retinal image. The results obtained in Part I rule out the possibility of attributing the results to changes in area. Whether the variation of threshold can be attributed to changes in curvature of the border independent of changes in area, has to be determined by the following experiment.

The Effect of Varying the Curvature of a Wavy Border upon Its Visibility

The wavy border test object illustrated in Fig. 6B provides a more satisfactory method of demonstrating the dependence of the visibility of a border upon its curvature than the disk type of test object because area as a variable is eliminated.

The experiment was carried out by starting with the finest wave pattern. This particular wave pattern has two thresholds; one at which the whole contour is perceived as a straight line, and a second threshold at which the individual waves are visible. The operator set the value of ΔB at a sub-threshold level, and the subject fixated the fixation point on the left and gradually increased ΔB till the straight border was visible, and then further increased it until the individual waves were visible. The subject then removed his hand from the control knob and the operator reset the value of ΔB to a sub-threshold level chosen at random. Five such sets of readings were made. The next coarser wave pattern was then substituted for the original pattern. With this pattern and with all coarser wave patterns, only one threshold was obtained; that is, the individual waves were visible whenever the contour on the left was visible at all. With this pattern, five threshold measurements were made for the wavy edge going from invisibility to visibility.

The same procedure was followed for the next coarser wave pattern and for the coarsest. After five settings were made for the wavy edge of the coarsest pattern, the subject fixated the point on the right and five settings were made for the threshold for the straight edge. The whole experiment was then repeated backwards. In Table V ascending order refers to the first half of the experiment, and descending order to the second half.

TABLE V. Contrast thresholds ($\Delta B/B$) for wavy borders constructed of quarter-circles of different curvature.

Subject	Order of test objects	Radius of curvature of arcs 5.2′ 1st	5.2′ 2nd	10.4′	20.7′	41.5′	Straight border
G.A.F.	Ascending	.0723	.1480	.0937	.1042	.0850	.0665
	Descending	.0665	.1300	.0909	.0986	.0751	.0658
	Average	.0694	.1390	.0923	.1014	.0800	.0662
J.O.F.	Ascending	.0453	.0845	.0576	.0566	.0584	.0435
	Descending	.0484	.0948	.0545	.0682	.0525	.0479
	Average	.0469	.0897	.0561	.0624	.0555	.0457

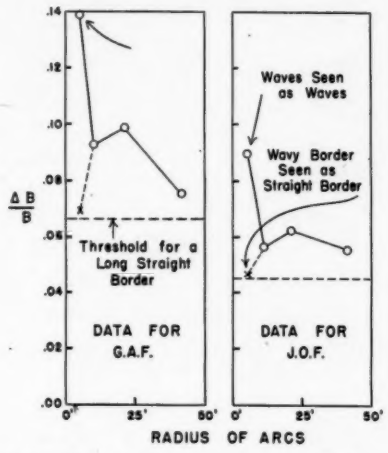

FIG. 13. The effect of varying the curvature of the waves in a wavy border upon its visibility.

The results for the two halves of the experiment showed no significant differences and were averaged. They are plotted in Fig. 13. In the case of the finest wave pattern, the cross represents the threshold at which the wavy edge is seen as a straight line, and the open circle represents the threshold at which the waves are visible as waves.

The fact that at the threshold level a wavy border will be transformed by the visual mechanism to a straight line constitutes a striking demonstration that the mechanism receiving and transmitting the impression of a border is much more adapted for long straight borders than for curved ones. This tendency for the wavy edge to be transformed to a straight line actually interferes with the perception of the waves as waves, because the threshold for a wavy edge is higher than for a disk having a border of the same curvature. The fact that the threshold for the 8-cycle pattern is lower than for the 4-cycle pattern, indicates that the forces that tend to transform the wave pattern into a straight line are not able to break down the wave pattern, but do aid in lowering the threshold.

Another way of describing the results is to say that the arrangement of the arcs so that their chords form a continuous straight line tends to lower threshold, and the shorter the chords the greater the effect until finally the wavy edge is transformed into a straight edge. On the other hand, curvature tends to raise the threshold and

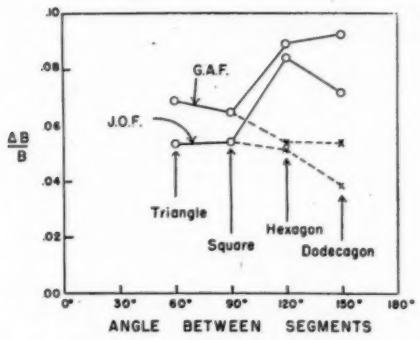

FIG. 14. Contrast thresholds for several regular polygons having sides of the same length.

the shorter the radius of curvature the greater the effect. In the case of the patterns having as few cycles as 4 or less, curvature dominates the situation, but for patterns having a greater number of cycles than 4, the other factor begins to outweight curvature.

It is pointless to try to carry the experiment to arcs of shorter radius than the shortest one in the experiment, because as soon as the radius of the arc approaches a magnitude comparable to the radius of a blur circle of a point source, the experiment loses its significance.

PART IV. THE RELATION OF THE ANGLE BETWEEN THE SEGMENTS OF A BORDER TO ITS VISIBILITY

Stimulus Patterns

This part of the paper is concerned with segmented borders of two types. One type is the segmented border of a regular polygon. In order to investigate this type of border, a stimulus pattern like that shown in Fig. 1A was used, in which regular polygons (a triangle, a square, a hexagon, and a dodecagon) were substituted for the rectangle. The length of the sides of all of the polygons was kept constant at 20.7′, and the variable from polygon to polygon to which importance was attached was the angle between the segments.

The saw-tooth pattern in Fig. 1B represents the second type of segmented border. The length of the segments (20.7′) was kept constant and the angle between the segments was varied as follows: 30°, 60°, 120°, and 150°. The straight edge on the right side of the test object can also be considered to be made up of segments with an angle of 180° between them.

Apparatus

The apparatus is the same as that described in Part I, except that special slides were used to produce the various polygon-shaped test objects and saw-tooth patterns of different pitch.

Procedure and Results

The same procedure was followed for the polygons as for the disks. However, in the case of the hexagon and dodecagon, the subjects encountered two thresholds, a low one at which the pattern was seen as a disk, and a higher one at which the individual straight sides of the polygon could be seen. Both thresholds were recorded. The results are given in Table VI and are plotted in Fig. 14.

The results for the saw-tooth patterns of various pitch are given in Table VII and plotted in Fig. 15.

Conclusions

The general conclusion can be drawn in the case of polygons that the size of the angle between the segments does not influence the visibility until the number of sides is increased to the point where the polygon is virtually a disk, and at the threshold level becomes visually transformed to a disk.

In the case of the saw-tooth pattern, the size of the angle between the segments does not influence the visibility until the angle approaches 180° so that the segments form virtually a straight line, and become visually transformed to a straight line at the threshold level.

TABLE VI. Contrast thresholds ($\Delta B/B$) for various regular polygons having sides of the same length (20.7′).

Subject	Order of test objects	Angle between segments 60°	90°	120° 1st	120° 2nd	150° 1st	150° 2nd
G.A.F.	Descending	.0645	.0696	.0527	.0903	.0488	.0903
	Ascending	.0730	.0598	.0543	.0893	.0589	.0950
	Average	.0688	.0647	.0534	.0898	.0539	.0927
J.O.F.	Descending	.0485	.0550	.0484	.0798	.0344	.0656
	Ascending	.0576	.0534	.0542	.0899	.0434	.0780
	Average	.0531	.0542	.0513	.0849	.0389	.0718

PART V. SUMMARY

By using rectangular test objects for measuring the contrast threshold, it has been possible to demonstrate that varying the width of a rectangle does not affect the threshold till the width is reduced to a size comparable to the diameter of the blur circle. The conclusion is drawn that the area of the retina covered by the test object (number of photo-receptors exposed) does not affect the threshold.

Varying the length of a rectangle does affect the threshold, and this is attributed to the length of the borders on the long sides of the rectangle. A saw-tooth pattern has also been used and it was found that varying the length of the segments affects the threshold. This affords further proof that the length of a border affects its visibility. The fact that the short and long sides of a rectangle become visible at the same time when the threshold is being measured, means that the visibility of the short sides as well as that of the long sides is affected by the length of the long borders. Whatever establishes visibility for the long borders also establishes it for the short ones.

On the other hand, the fact that varying length of the short borders does not affect the threshold, indicates that the length of the short borders does not affect the visibility of the long ones.

By using regular polygons having different numbers of sides, and by using saw-tooth borders of different pitch, in which the straight segments are all the same length, it is possible to demonstrate that the angle between the segments has no effect upon the threshold until the angle is so large that the segmented border is transformed visually to a continuous border. This finding is important for the formulation of a theory of mechanism involved in the perception of a contrast border, because so far as the threshold is concerned it means that consideration can be limited to the case of segments joining each other at right angles. By using disks and wavy borders, it has been possible to demonstrate that the curvature of a border affects its visibility such that the higher the curvature the higher becomes the threshold. If the pitch of a right-angled saw-tooth border or a wavy border is fine enough, the border appears as a straight edge at the threshold level, which indicated that the contrast border mechanisms of the eye and brain are better adapted for a long straight border than for a saw-tooth or wavy border.

TABLE VII. Contrast thresholds ($\Delta B/B$) for saw-tooth borders differing in pitch but having segments of the same length (20.7′).

Subject	Order of test objects	Angle between segments 30°	60°	90°	120°	150° 1st	150° 2nd	180° Straight border
G.A.F.	Descending	.0687	.0711	.0773	.0770	.0508	—	.0598
	Ascending	.0870	.0728	.0778	.0785	.0715	—	.0561
	Average	.0779	.0720	.0776	.0778	.0612		.0580
J.O.F.	Descending	.0465	.0437	.0508	.0502	.0435	.0809	.0362
	Ascending	.0492	.0454	.0526	.0496	.0435	.0688	.0420
	Average	.0479	.0446	.0517	.0499	.0435	.0749	.0391

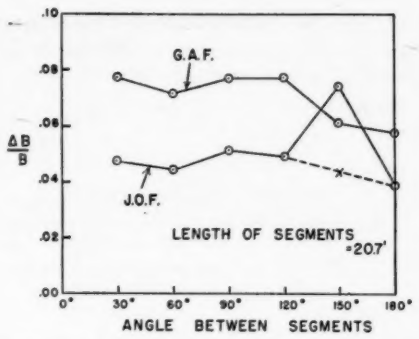

FIG. 15. The effect of varying the angle between the segments in a saw-tooth border upon its visibility.

For those who are interested in the nervous mechanisms connected with the eye, the investigations have pointed up the following problems.

1. What determines the brightness difference required for a threshold?
2. Why does the length of a border affect its visibility?
3. Why do the long sides of a rectangle affect the visibility of the short sides?
4. Why does the curvature of a border affect its visibility?
5. Why does a wavy or saw-tooth border having a fine pitch become transformed to a continuous straight border at the threshold level, and why does a regular polygon whose border is made up of a large number of small straight segments appear as a disk at the threshold level?

A Theoretical Development of Exact Solution of Reflectance of Multiple Layer Optical Coatings

DORIS L. CABALLERO
Bell and Howell Company, Chicago, Illinois
(Received October 14, 1946)

A simple method is developed for calculation of reflectance of homogeneous multiple layer coatings for the case of normal incidence. The derivation is based upon a vectorial representation of the reflectance of a single or more coating layers, found by a summation of the infinite series of multiple reflections within each layer. Presence of absorption and dispersion may be accounted for in the formulas.

SUMMARY

TO facilitate the design of multiple layer optical coatings, it has become necessary to develop a method which would enable quick and accurate prediction of reflectance characteristics. Previous theoretical discussions based on consideration of multiple reflections have been incomplete or too complex for practical application. An exact theoretical treatment of the problem has been presented by considering the situation as a boundary problem in electromagnetic theory.[1-2] The development has been extended to include optical coatings consisting of as many as five layers. However, the formulas are quite complex and computations require a considerable amount of time.

It has been found that an analysis involving vectorial representation of the reflectance of a single or more coating layers, found by a summation of the series of inter-reflections within each layer, requires considerably less specialized knowledge and results in a simpler method of derivation. The formulas may be solved either analytically or graphically and furnish an exact solution of coating reflectance for a given wavelength of light at normal incidence. The method is applicable to coatings consisting of any number of layers and is useful in prediction of reflectance or transmission characteristics of multiple layer optical coatings. Computation by this method requires less time than with the exact methods previously developed.

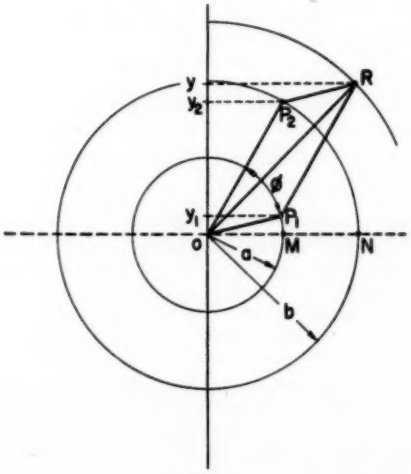

FIG. 1.

THEORETICAL CONSIDERATIONS

As a brief introduction to the method employed in this analysis, the following well-known theoretical considerations are presented.

Displacements in two trains of waves vibrating with simple harmonic motion of equal period, T, but with unequal amplitudes, a and b, and different phases (due to path difference), may be represented by the equations[3]

$$y_1 = a \sin 2\pi(t/T - x_1/\lambda),$$
$$y_2 = b \sin 2\pi(t/T - x_2/\lambda),$$

where y_1 and y_2 are the displacements at distances x_1 and x_2 from the source at any time, t. The difference in phase is, then,

$$(2\pi/\lambda)(x_2 - x_1) = 2\pi/\lambda \cdot \text{path difference}. \quad (1)$$

[1] R. L. Mooney, J. Opt. Soc. Am. **35**, 574–583 (1945).
[2] R. L. Mooney, J. Opt. Soc. Am. **36**, 256–260 (1946).
[3] J. K. Robertson, *Introduction to Physical Optics* (D. Van Nostrand Company, Inc., New York, 1941), pp. 46–47, pp. 36–37.

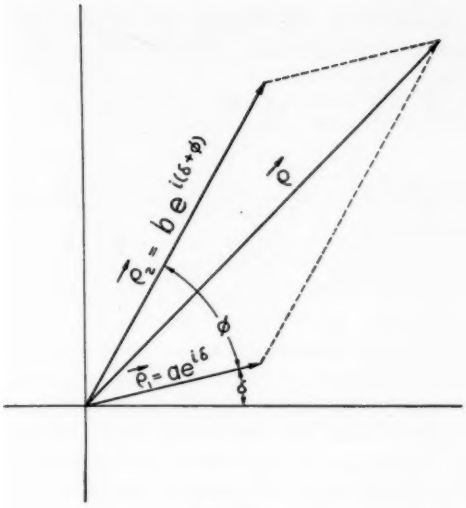

FIG. 2.

The amplitude and phase of the resultant of the two wave trains may be found graphically as shown in Fig. 1, where φ equals the difference in phase. To represent this situation in a simpler form, the two amplitudes OP_1 and OP_2 and their phase condition may be expressed by two-dimensional vectors as

$$\varrho_1 = ae^{i\delta},$$
$$\varrho_2 = be^{i(\delta+\varphi)},$$

where a and b are magnitudes, δ is $2\pi t/T$, and φ is the phase difference. The graphical solution is as shown in Fig. 2 and illustrates the method of representation of waves to be used in the forthcoming derivations.

If a wave is normally incident upon a thin film (although shown at an angle in Fig. 3 for convenience in representation), the primary amplitudes, ρ_1 and ρ_2, reflected from the interfaces of the film may be found by application of Fresnel's equations[4] for normal incidence. For the first interface,

$$\rho_1 = [(n_1-n_2)/(n_1+n_2)]\rho_I,$$

where ρ_I is the incident wave amplitude, ρ_1 the reflected amplitude, and n_1, n_2, n_3, are the indices of refraction of the medium of incidence, of the film, and of the material upon which the film is deposited, respectively. If n_2 is greater than n_1 it is apparent that the quantity is negative, representing the 180° phase change due to reflection from a higher index medium. Letting the fraction $(n_1-n_2)/(n_1+n_2)=A$, it follows that the amplitude,

$$\rho_1 = A\rho_I.$$

Next, to find ρ_2, the fraction of incident amplitude transmitted through the first interface is[5]

$$\rho_T' = \left\{\left[1-\left(\frac{n_1-n_2}{n_1+n_2}\right)^2\right]\frac{n_1}{n_2}\right\}^{\frac{1}{2}}\cdot\rho_I$$
$$= [(1-A^2)(n_1/n_2)]^{\frac{1}{2}}\cdot\rho_I.$$

This amplitude, ρ_T' is incident upon the second interface and is partially transmitted and partially reflected. The amplitude reflected is,

$$\rho_R' = [(n_2-n_3)/(n_2+n_3)]\cdot[(1-A^2)(n_1/n_2)]^{\frac{1}{2}}\cdot\rho_I$$
$$= B\cdot[(1-A^2)(n_1/n_2)]^{\frac{1}{2}}\cdot\rho_I,$$

letting $B=(n_2-n_3)/(n_2+n_3)$. Then this amplitude is transmitted back through the first interface, so that the final value of ρ_2 is

$$\rho_2 = [(1-A^2)(n_1/n_2)]^{\frac{1}{2}}$$
$$\cdot[(1-A^2)(n_2/n_1)]^{\frac{1}{2}}\cdot B\cdot\rho_I = (1-A^2)\cdot B\cdot\rho_I.$$

Assuming unit incident intensity, $\rho_I=1$, and the equations for the amplitude of the two waves are,

$$\rho_1 = A,$$
$$\rho_2 = (1-A^2)B.$$

It has been shown that the path difference of waves striking a film at normal incidence is $2nt$,[6] where t is the film thickness, and n is the refractive index of the film. Thus the phase

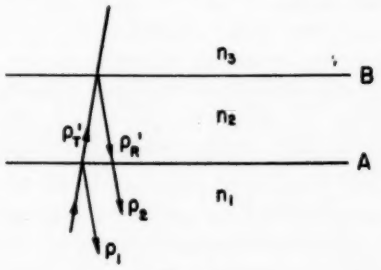

FIG. 3.

[4] R. W. Wood, *Physical Optics* (The Macmillan Company, New York, 1934), pp. 410–411.

[5] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), pp. 494–497.

[6] J. K. Robertson, *Introduction to Physical Optics* (D. Van Nostrand Company, Inc., New York, 1941), pp. 226–230.

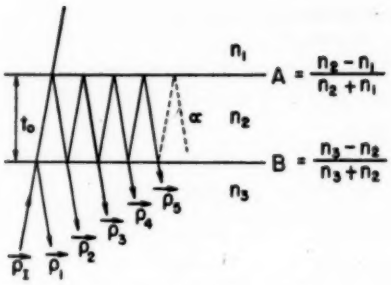

FIG. 4.

difference (from Eq. (1)) is,

$$(2\pi/\lambda)\cdot 2nt = 4\pi nt/\lambda$$

and since it is customary to speak of an optical film thickness, t_0, measured in terms of wave-lengths, where $t_0 = t/\lambda/n$, the expression is simplified to $4\pi t_0$.

Letting $\alpha = 4\pi t_0$, the amplitude and phase of the primary reflections may be expressed by,

$$\varrho_1 = Ae^{i0},$$
$$\varrho_2 = (1-A^2)Be^{i\alpha}.$$

The quantities A and B will be referred to as the reflection coefficients of the interfaces.

DERIVATION OF EQUATIONS FOR EXACT REFLECTANCE OF COATINGS

1. Single Layer Coating

Thus far only the primary reflections from the interfaces of the film or coating have been considered. However, as can be seen from Fig. 4, successive inter-reflections between the interfaces of the coating result in an infinite number of reflected waves of decreasing amplitude and successive phase difference of α. Letting n_1, n_2, n_3, represent the indices of refraction of the glass, coating, and medium of incidence, respectively, the series of reflections can be expressed in accordance with the preceding theoretical discussion.

$$\varrho_1 = Be^{i0},$$
$$\varrho_2 = (1-B^2)Ae^{i\alpha},$$
$$\varrho_3 = (1-B^2)A^2(-B)e^{2i\alpha},$$
$$\varrho_4 = (1-B^2)A^3(-B)^2e^{3i\alpha},$$
$$\cdot \quad \cdot \quad \cdot$$
$$\cdot \quad \cdot \quad \cdot$$
$$\varrho_j = (1-B^2)A^{j-1}(-B)^{j-2}e^{(j-1)i\alpha}.$$

Exclusive of ϱ_1, the vectors form a convergent geometric series, the sum of which may be shown to be,[7]

$$S = \frac{(1-B^2)Ae^{i\alpha}}{1-A(-B)e^{i\alpha}}.$$

Adding ϱ_1 to the series, the resultant of the series of vectors is,

$$\varrho_A = B + \frac{(1-B^2)Ae^{i\alpha}}{1-A(-B)e^{i\alpha}} = \frac{B+Ae^{i\alpha}}{1+ABe^{i\alpha}}.$$

This is the final equation for computations of reflectance amplitude of a single layer coating, and may be solved either graphically or analytically. Unless the calculations are continued for a two-layer coating, it is not necessary to designate the phase condition, or direction, of the resultant, since the magnitude alone will indicate reflectance amplitude. By squaring this value, the reflectance intensity is obtained. It is customary to consider the incident intensity as 100 percent and to plot values of percent of incident intensity reflected by the coating against the wave-length. The reflected intensity, of course, must be calculated for various wave-lengths, since phase retardation due to traversal of a coating is a function of λ.

2. Two-Layer Coating

Figure 5 represents a two-layer coating of indices n_2 and n_3 on glass of index n_1. The re-

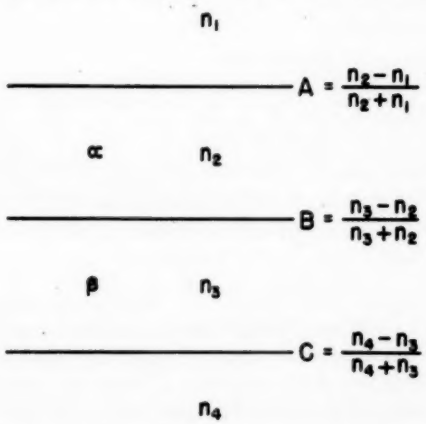

FIG. 5.

[7] C. E. Love, *Differential and Integral Calculus* (The Macmillan Company, New York, 1934), p. 254.

fractive index of the medium of incidence is n_4. The interface reflection coefficients are A, B, C, and the second layer is of an optical thickness such as to produce a phase change of β in light of given wave-length traversing twice the coating thickness.

In the preceding section, a method of computing exact reflectance amplitude of a single layer was outlined. Therefore, the reflectance of the first coating layer (index n_2) can be found. Let this resulting amplitude, $\varrho_A=(B+Ae^{i\alpha})/(1+ABe^{i\alpha})$, equal $\mathbf{Z}$, representing the effect of the first layer upon any wave incident upon that layer. By employing $\mathbf{Z}$ as an interface reflection coefficient, multiple reflections are considered without use of hopelessly complicated formulas. The problem then is similar to that of a single layer coating as may be seen by Fig. 6. The series of reflections are as follows:

$$\varrho_1=Ce^{i0},$$
$$\varrho_2=(1-C^2)\mathbf{Z}e^{i\beta},$$
$$\varrho_3=(1-C^2)\mathbf{Z}^2(-C)e^{2i\beta},$$
$$\varrho_4=(1-C^2)\mathbf{Z}^3(-C)^2e^{3i\beta},$$
$$\vdots$$
$$\varrho_j=(1-C^2)\mathbf{Z}^{j-1}(-C)^{j-2}e^{(j-1)i\beta}.$$

The resultant reflectance amplitude is,

$$\varrho_B=C+\frac{(1-C^2)\mathbf{Z}e^{i\beta}}{1-(-C)\mathbf{Z}e^{i\beta}}=\frac{C+\mathbf{Z}e^{i\beta}}{1+C\mathbf{Z}e^{i\beta}}=\mathbf{Y}.$$

Solution of this formula may be found graphically or analytically.

3. Multiple Layer Coating

Proceeding one step further, the reflectance of a three-layer coating, Fig. 7, is determined by

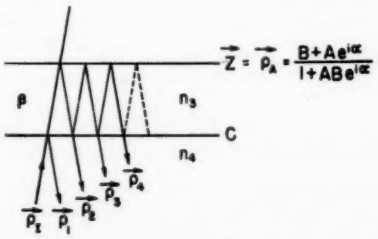

FIG. 6.

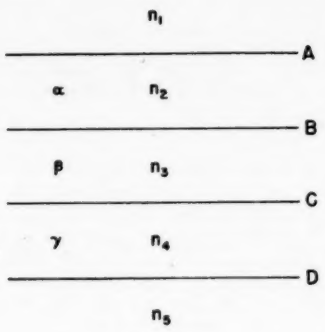

FIG. 7.

considering the reflectance of the first two layers (indices n_2 and n_3) as a vector $\mathbf{Y}$ found as shown in the preceding section. n_4 is the index of the third layer and n_5 is that of the medium of incidence. γ is the phase change due to traversal of the third layer. Again the problem is reduced to that of a single layer coating, and the sum of the infinite series of vectors may be obtained. The resultant is,

$$\varrho_C=\frac{D+\mathbf{Y}e^{i\gamma}}{1+D\mathbf{Y}e^{i\gamma}}=\mathbf{W}.$$

This method may be continued for solution of a four-layer coating by treating the first three layers as an interface of reflection coefficient $\mathbf{W}$, calculated as just shown, and resolving the infinite series of reflections between it and the coating-air interface. Thus, by a successive vectorial representation of the reflectance of one, two, or more layers, the resultant reflectance of any multiple layer optical coating may be found.

ABSORPTION

Absorption has not ordinarily been included in coating computations since its effect has been found negligible due to the extreme thinness of the coating layers. However, an investigation may be found necessary if discrepancies arise between theoretical and actual data when predicting coating reflectance characteristics. If the absorption of a coating material, expressed as a fraction of incident intensity per centimeter of the material, is known for a given wave-length, it may be included in the calculations in the following manner:

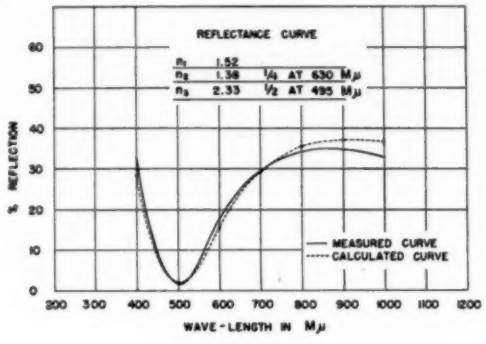

FIG. 8.

Given a single layer coating of thickness t in centimeters with an absorption measured to be a fraction, k, of incident intensity per centimeter of thickness for a given wave-length λ, each intensity traversing the coating thickness, may be multiplied by kt to obtain the amount absorbed by the coating. In this case, however, it is more desirable to use a fractional coefficient $(1-kt)=N$, thus obtaining directly the remaining intensity after absorption has taken place. Since intensity is proportional to wave amplitude, each amplitude traversing the coating thickness (back and forth) may be multiplied by $\sqrt{N}\cdot\sqrt{N}=N$. The solution for a single layer coating reflectance amplitude will then be,

$$\rho_A=\frac{B+ANe^{i\alpha}}{1+ABNe^{i\alpha}}.$$

In the same manner, the solution of a two-layer coating with absorption coefficients N_1 and N_2 of the separate layers, will change to read,

$$\mathbf{Y}=\frac{C+\mathbf{Z}N_2e^{i\beta}}{1+C\mathbf{Z}N_2e^{i\beta}}$$

where

$$\mathbf{Z}=\frac{B+AN_1e^{i\alpha}}{1+ABN_1e^{i\alpha}}.$$

It follows that the formula for a three-layer coating will be,

$$\mathbf{W}=\frac{D+\mathbf{Y}N_3e^{i\gamma}}{1+D\mathbf{Y}N_3e^{i\gamma}}.$$

CONCLUSION

An analysis has been presented which makes it possible, by means of comparatively simple formulas, to predict exactly the reflectance of optical coatings of any number of layers.

Calculation for absorption is possible without unduly increasing the complexity of the formulas.

Presence of dispersion in a coating material may be accounted for by substitution of the correct indices of refraction at various wave-lengths for which reflectance is calculated.

Figure 8 shows reflectance curves of a two-layer coating, one obtained by actual measurement and one calculated by the method outlined in this report. Discrepancies between the two curves are caused by the absorption of the glass, coating absorption, errors in assumed indices and coating thickness, and inaccuracies in measured reflectances. Calculations using previous exact theoretical equations result in a reflectance curve identical with that shown in Fig. 8.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

Effect of a Constant Current on Vision

S. V. KRAVKOV AND L. P. GALOCHKINA
Helmholtz Central Ophthalmological Institute, Moscow, USSR

(Received September 9, 1946)

Experiments described in this paper demonstrate that a weak constant current when running through the eye ball may noticeably alter the sensitivity of rod vision as well as that of cone vision. The direction of these changes of sensitivity depends upon which pole is placed on the eye ball. In the case of cone vision this direction is also influenced by the wave-length of the light stimulus. The authors advance an hypothesis explaining the action of the constant current on the eye by changes in the relative concentration of Ca and K ions in the neighborhood of poles. Special experiments performed with Ca and K iontophoresis have confirmed this interpretation.

THE possibility of inducing visual sensation by means of switching on and off an electrical current running through the eye ball has been known since the middle of the eighteenth century, and has been the object of numerous investigations the purpose of which was to elucidate the so called electrical sensitivity of the eye.[1]

When carrying out the present investigation, we were interested in another problem concerning the action of the electrical current on the eye, namely in changes, and in the character of changes, of eye sensitivity, if any, to adequate light stimuli arising under the influence of the current.

Data contained in the literature on changes of adequate sensitivity occurring as a result of electrical excitation are rather scarce and mostly of a qualitative character.

Thus, the investigations of Ritter,[2] Schelske,[3] Aubert,[4] Schliephake,[5] and Helmholtz[6] revealed that when the anode is placed on the eye and the cathode on the occiput or on the hand, which means an ascending current, its switching on involves a lighter visual field, which at the same time acquires a lilac-blue tinge. Conversely, when the cathode is placed on the eye (descending current), connecting the current brings about a darkening of the visual field and a reddish-yellowish tinge. The descending current in the moment of connection and during some period of its subsequent passage through the eye calls forth greater distinctness of visible objects, while the ascending current, in conformity with the findings of the above mentioned authors, causes lesser distinctness of these objects. The authors are inclined to interpret this phenomenon as a result of changes arising under the effect of the electrical current in the so called intrinsic light of the eye.[3,6]

Some experiments supporting the fact that electrical stimulation of the eye is able to bring changes in its adequate sensitivity were also performed by Kravkov, Semenovskaja, and Bogoslovsky in 1935.[7]

METHOD

Our chief experiments consisted in determination of absolute thresholds for peripheral rod vision and of those for foveal cone vision. The determination of thresholds was performed monocularly. Changes in the intensity of the light stimulus were effected by inserting and withdrawing a neutral photo-wedge. We devised an adaptometer for the special purpose of measuring the thresholds of peripheral vision. Those of foveal vision were determined by means of a

[1] A. I. Bogoslovsky, "The problem of the electrical sensitivity of the eye," *Problems of Physiological Optics* (Moscow-Leningrad, 1944), Vol. 2, p. 136. In Russian, with English summaries.

[2] J. Ritter, "Versuche und Bemerkungen über den Galvanismus der voltaischen Battarie," Ann. d. Physik **7**, 431 (1801).

[3] R. Schelske, "Zur Farbenempfindung," Arch. f. Ophthalm. **9**, 39 (1863).

[4] H. Aubert, *Physiologie der Netzhaut* (Breslau, 1865), p. 345.

[5] H. Schliephake, "Beiträge zur Kenntniss der Einwirkung des galvanischen Stromes auf das menschliche Auge," Arch. ges. Physiol. **8**, 567 (1873).

[6] H. v. Helmholtz, *Handbuch d. physiologischen Optik* (Hamburg-Leipzig, 1941), Vol. 2, p. 16.

[7] S. V. Kravkov, E. N. Semenovskaja, and A. I. Bogoslovsky, *Visual Sensations and Perceptions* (Moscow, 1935), p. 176. (In Russian.)

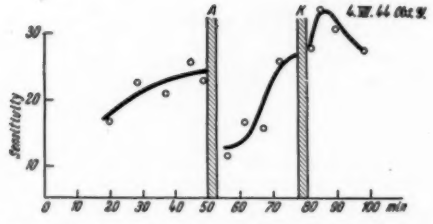

FIG. 1. Changes of rod sensitivity following the passage of a constant current through the eye, with the anode on the eye ball (A) or the cathode on the eye ball (K). Current intensity 0.02 ma.

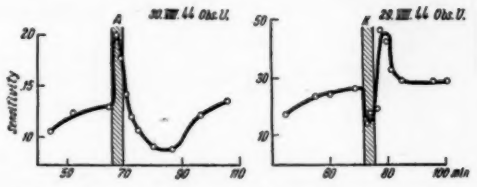

FIG. 2. Changes of rod sensitivity during the passage of a constant current through the eye and after the current has been switched off. A—anode on the eye. Intensity—0.02 ma.

monchromator. Experiments were carried out under conditions of complete dark adaptation.

The eye was subjected to stimulation by a constant electrical current after a certain stable level of sensitivity had been attained in the process of dark adaptation and the experiment was continued for more than 30 minutes; in the course of this time additional determinations of the threshold were performed. The strength of the current varied in different series of experiments from 0.02 to 0.5 ma; its duration from 3 to 20 minutes. Silver plates with gauze moistened with saline served as electrodes. In various series of experiments the active electrode was placed either on the temple near the temporal angle of the eye or on the lids of the closed eye; more frequently on the lids of an open eye. The electrode had the form of an oval plate with an orifice in the middle. The indifferent electrode was placed on the posterior part of the individual's neck or, more commonly, he took it in the hand of the side opposite to the tested eye. The threshold was measured in values characterising the transmission in the corresponding sites of the photo-wedge; sensitivity was determined as a value inversely proportional to that of the threshold. The value of sensitivity present prior to electrical stimulation, adopted in all the figures below, was 100.

RESULTS

A. Rod Vision

These series of our experiments were carried out on 13 individuals. They all proved that stimulation of the eye with a constant current involves obvious changes in the level of peripheral sensitivity of rod vision. Both a very weak current (0.02 ma, duration 3 minutes) and a stronger one (0.2 ma, duration 10 min.) were used in these series. The changes of sensitivity arising under the effect of the current proved quite opposite in character depending upon which pole of the current (the anode or the cathode) was on the eye. In the case of the anode on the eye the effect was always a subsequent decrease of rod sensitivity. Conversely the cathode raised the subsequent sensitivity to light. This polar effect is clearly shown on Fig. 1, where the findings of an experiment are given; in this test the electrical stimulus was at first used with the anode on the eye (A) and the direction of the current changed afterwards in the same experiment (K).

During the passage of the current through the eye, the light sensitivity of rod vision displayed obvious changes, which however were the reverse of those recorded in the after effect. Thus, the descendent current (with the cathode on the eye) lowered sensitivity, while the ascendent one, with the anode on the eye, increased it. Figure 2, where the typical findings of two experiments are pictured, illustrates this statement.

In control experiments the electrodes were placed on the temples of the tested individuals on the external angles of their eyes. Under these conditions, rod sensitivity was raised in the eye near the anode, while conversely the eye near the cathode showed lowered sensitivity.

B. Cone Vision

In these series of experiments the determination of the threshold of visibility of various monochromatic spectral rays, perceived only in the foveal area of the retina, was performed using the extinction method. These tests were carried out on 5 individuals. The results showed no discrepancy and were as follows: under condi-

ions of an electrotonus, e.g., in the passage of an ascendent current through the eye (the anode on the eye)—the cone sensitivity to green-blue spectral rays proved augmented while it was diminished to orange-red ones. Conversely, during the passage of the descendent current (the cathode on the eye) under conditions of cathelectrotonus, decrease of sensitivity towards blue-green rays and its increase towards orange-red ones was invariably recorded. Three regions of the spectrum—the two terminal ones and the field of yellow (near 570 millimicrons) appeared as neutral; in other words, cone sensitivity towards these rays failed to show changes in either direction under the action of an electrical current (intensity from 0.02 to 0.2 ma and duration from 3 to 15 minutes).

A characteristic picture of an individual experiment in which different directions of the current were used (anode on the eye, or cathode on the eye) together with various monochromatic stimuli, is given on Fig. 3 with typical data of four experiments performed on the same individual.

The after-effect of the current commonly showed changes in cone sensitivity quite opposite to those which were recorded during electric stimulation.

Experiments on changes of cone sensitivity occurring under the influence of a constant current were carried out using monochromatic rays with a wave-length of 425, 470, 520, 570, 600, 620, 650, and 700 millimicrons as light stimuli.

Curves, picturing the results of experiments performed on 5 individuals, subjected to the action of a constant current of 0.02 ma, are given on Fig. 4. The length of waves in millimicrons is put on the abscissa, and the ratio S_1/S (S-value of cone sensitivity prior to the passage of the current through the eye and S_1-value of cone sensitivity at the end of the third minute of

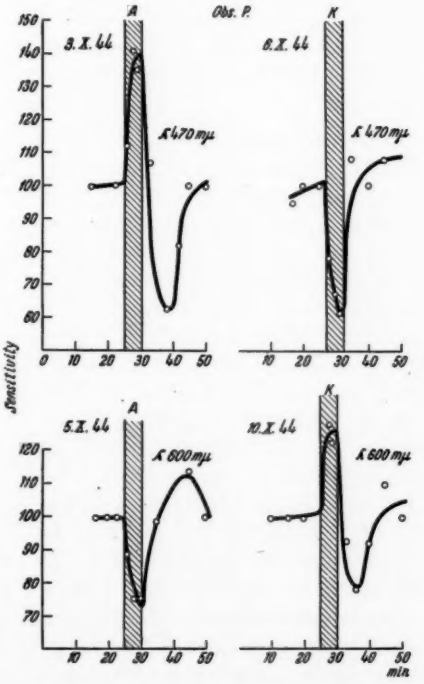

FIG. 3. Changes of cone sensitivity to rays of various wave length under the influence of a constant current passing through the eye. A—anode on the eye. K—cathode on the eye. Intensity—0.02 ma.

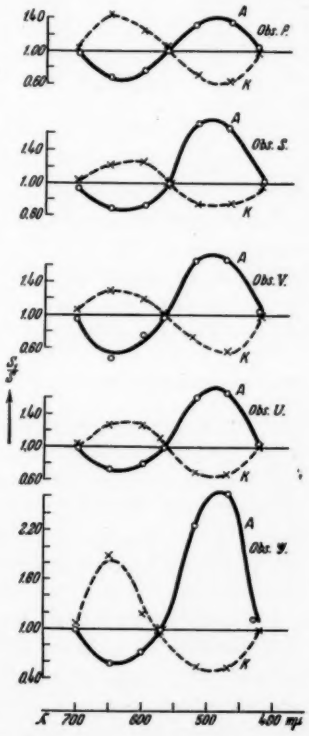

FIG. 4. Changes of cone sensitivity during the passage of a constant current. A—anode on the eye. K—cathode of the eye. Intensity—0.02 ma. Data obtained from the investigation of 5 individuals.

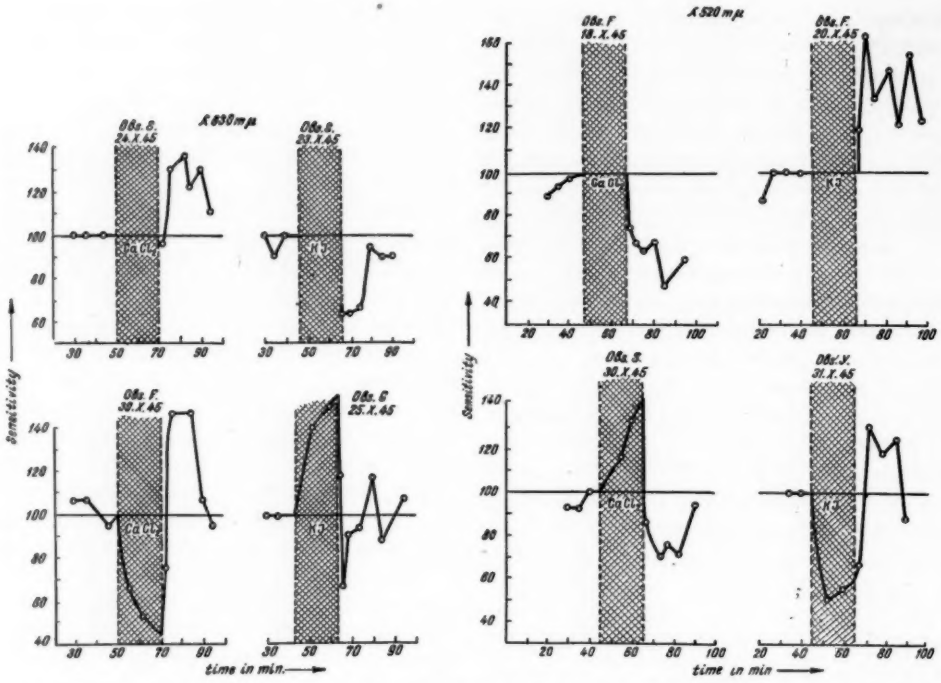

FIG. 5. Effect of Ca and K iontophoresis on cone sensitivity of the eye to red (630 m) and green (520 m) color. Intensity—0.5 ma.

current action) is marked on the ordinate. Values of sensitivity obtained under conditions of anelectrotonus are represented by continuous lines, the letter *A* and small circles; those obtained under conditions of cathelectrotonus—by dotted lines, the letter *K*, and crosses.

DISCUSSION

What, then, are the processes promoted by a constant current in our visual apparatus which are responsible for the above described changes in cone and rod sensitivity? We are inclined to suppose that these changes can be accounted for by changes in the ratio of concentration of potassium and calcium ions brought about by the passage of a constant current. Monovalent potassium ions shift towards the cathode with greater velocity than bivalent calcium ions; therefore, the ratio potassium/calcium ions is increased near the cathode and diminished near the anode.

Loeb's investigations[8] demonstrated the significance of the presence of potassium and calcium ions in live tissues as regards some of their properties, particularly excitability. It is, moreover, well known that the physiological action of potassium and calcium ions is mutually antagonistic. Potassium provokes tissue edema while calcium contributes to diminishment of swelling. Potassium raises the dispersion of colloidal substances, whereas calcium lowers it. Prevalence of potassium ions leads to alkalosis; that of calcium ions, on the contrary, calls forth acidosis. These properties are connected with capillary-electrical phenomena in tissues. Hence changes of potassium-calcium balance involve deep alterations of cellular life conditions. No wonder, therefore, that changes of the potassium/calcium ratio, brought about by the action of a constant current, may be responsible for altered excitability of nervous visual cells as well. Reverse variation of cone sensitivity towards green-blue and orange-red spectral rays, as influenced by a constant current passing in the same direction, is suggestive of some contrasting chemico-physical

[8] J. Loeb, *Vorlesungen über die Dynamik der Lebenserscheinungen* (Leipzig, 1906).

features of those nervous apparatus which are apt to be stimulated by short and long wave spectral rays.

In this connection a law, established in physiology, mainly by the investigations of Zondek,[9] is of particular interest. In many cases the effect of potassium is similar to that of a stimulation of the parasympathetic vagus nerve, whereas the action of calcium resembles that of excitation of the sympathetic nerve.

The antagonistic action of the sympathetic and parasympathetic nervous system presents only a manifestation of changes in the function of cells which depend upon the prevalence of either calcium or potassium ions (Dresel[10]).

If we take into account all that has been said above, we must admit that excitability in some of our apparatus of vision is augmented under the influence of increased relative concentration of calcium ions in the environment (which occurs in the region of the anode) and, therefore, these parts of the visual apparatus are sympathetico-tropic—those perceiving green and blue rays in cone vision—whereas others take advantage of a relative increase of potassium ion concentration (near the cathode) and are to be considered as parasympathetico-tropic.[11]

As for our rod vision, its components, as regards reaction to electrical current, should be regarded as sympathetico-tropic.

EXPERIMENTS WITH IONTOPHORESIS

We carried out special tests bringing potassium and calcium ions into the eye (1 percent solution of potassium iodide and 2 percent solution of $CaCl^2$ by means of iontophoresis) for the purpose of checking our interpretation of the above described experiments. Pads made of gauze, moistened with either solution, were placed on the lids of a closed eye. It goes without saying that the anode was the active electrode placed on the eye. The indifferent electrode was put on the back of the neck of the individual under investigation. The intensity of the current used for iontophoresis was of 0.5 ma, its duration—20 min. Under conditions of dark adaptation and of monocular, strictly foveal, vision the threshold value of disappearance of visibility has been determined for monochromatic red rays as 630 mμ and as 520 mμ for green ones. In some experiments of these series the estimation of thresholds has been performed during iontophoresis. The active electrode placed on the closed eye was removed for a short time, during measurement of the threshold, and shifted towards the temple, near the external angle of the eye. These tests were checked on four individuals with normal color vision. The results demonstrated that the potassium and calcium ions, when brought into the eye, are antagonistic as regards their effect on cone sensitivity. These ions promote changes in sensitivity of the eye to red and green color in opposite directions. All this is illustrated on Fig. 5.

Our experiment with iontophoresis when placing a pad moistened with a potassium solution on one eye, and a pad moistened with a calcium solution on the other one, is quite demonstrative. We had thus the opportunity to see that under conditions of the same experiment and with the same pole of the current on both eyes, the cone sensitivity to green color in either eye (520 mμ) changed in the opposite direction, depending upon which ions were brought into the eye (see Fig. 6, where this phenomenon is illustrated).

Thus we believe that we have fully demon-

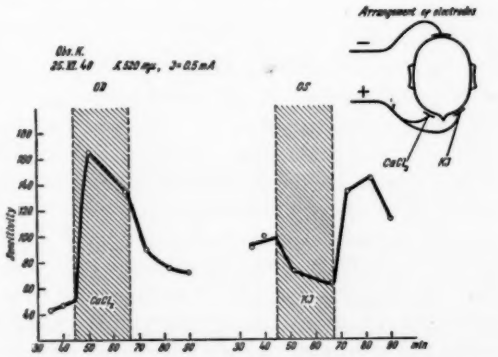

FIG. 6. Effect of Ca and K iontophoresis on cone sensitivity to green color (520 m), when both are brought into the right and left eye in the same experiment. Intensity —0.5 ma.

[9] S. Zondek, "Die Bedeutung des Antagonismus von Kalium und Calcium für die Physiologie und Pathologie. Klin. Wschr., No. 9, 382 (1923).—Identität von Nerv- und Ionenwirkung," Biochem. Zeits. **132**, 362 (1922).

[10] K. Dresel, *Erkrankungen des vegetativen Nervensystems* (Berlin-Wien, 1924).

[11] S. V. Kravkov, "Color vision and autonomic nervous system," J. Opt. Soc. Am. **32**, 335 (1941).

strated the role of potassium and calcium ions in the above described phenomena. The way in which this effect is actually accomplished is to be the object of further investigation.

SUMMARY

Cone and rod sensitivity show obvious changes under the influence of a constant current passing through the eye. This effect is produced by changes of relative concentration of potassium and calcium ions in the retina. The different response of various apparatuses of our vision to anelectrotonus and cathelectrotonus testify to their connection, either with the sympathetic or with the parasympathetic, portion of the autonomous nervous system.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

Apparatus for Determining Spherical Aberration of Convex Surface of Paraboloid of Revolution

BENJAMIN GOLDBERG AND NELS C. BENSON
Engineer Board, War Department, Fort Belvoir, Virginia

In the production of 60-inch anti-aircraft searchlight reflectors by the electroforming process, convex parabolic forms are used as molds. To avoid production of reflectors with excessive spherical aberration, it is necessary to know the inherent spherical aberration of each mold. To obtain this information, an apparatus has been devised to measure the spherical aberrations of the convex surface of a paraboloid of revolution. The optical system of the apparatus consists essentially of a collimator, a plane mirror, and a photographic plate used in the following manner. A narrow beam of light parallel to the optical axis of the mold to be tested is projected from the collimator onto a designated spot of a particular zone on the mold which is in a horizontal position. The beam of light is reflected from the mold as if it were coming from the virtual focus for the designated spot. The reflected light strikes the plane mirror, vertically placed in a predetermined position and is re-reflected so that the beam of light intersects the optical axis of the mold at a distance above the mold equal to the focal length for the designated spot in the mold being examined. The photographic plate is placed horizontally at a distance above the mold equal to the nominal focal length of the mold. The beam of light intercepts the photographic plate and produces a specified image. Eight such images are photographed on one plate at equally spaced intervals around the periphery of the zone being tested. A required number of zones are tested. Simple interpretation of the photographs yield the amount of aberration present. An important feature of this apparatus is that the interpretations of the photographs are identical to those obtained for testing concave parabolic reflectors. This apparatus aids reflector production in that it determines the abberation characteristics of the convex mold directly, thus eliminating the previous practice of making a reflector from a mold to determine the mold's accuracy indirectly.

I. PURPOSE

ANTI-AIRCRAFT searchlight reflectors are produced by an electroforming process. By this process, all-metal reflectors are electrolytically produced from convex paraboloidal molds. These molds whether manufactured of metal or glass are never geometrically true. Therefore, a small specified amount of aberration is permitted from a true paraboloidal surface. Molds having a larger amount of aberration than permitted by specifications are rejected. Originally, the only way to determine this fact was to make a reflector from the mold, test the reflector for its aberration characteristics, and so indirectly deduce the contour of the mold. This meant lost time and material if the reflector was faulty, since both the reflector and mold were rejected. To eliminate this difficulty an apparatus was devised to determine directly the aberration inherent in the convex paraboloidal form.

Before discussing the new apparatus, let us acquaint ourselves with the method used by the Engineer Board to determine aberration characteristics of 60-inch paraboloidal reflectors. Figure 1 shows the standard apparatus. The reflector is placed horizontally on the turntable and is properly centered. A narrow beam of light is projected from the collimator onto a designated spot of a particular zone on the reflector. The beam is parallel to the optical axis of the reflector.

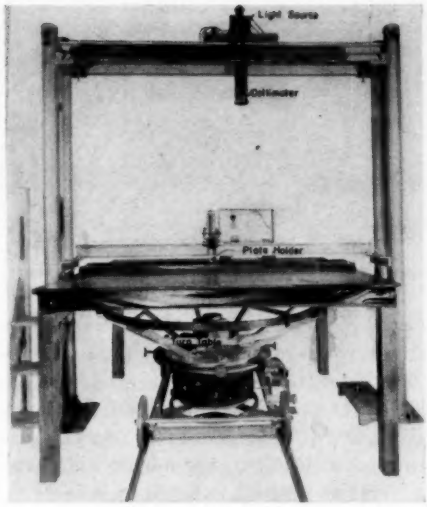

FIG. 1.

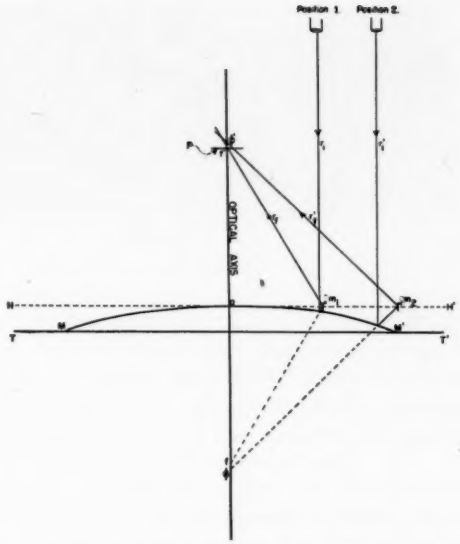

FIG. 2. Focal Point Test of Convex Surfaces.

TT'—turntable
MM'—negative mold
faf'—optical axis
HH'—horizontal line tangent to nose of mold
af—focal length of mold
$f\phi$—increase in focal length due to aberration in outer zone of mold
m_1 and m_2—plane mirrors
r_1 and r_1'—incident rays projected from collimator in position 1 and position 2
r_f and r_f'—reflected rays
P—photographic plate

The beam is reflected so that it intersects the optical axis at or near the nominal focal point of the reflector depending on the absence or presence of aberration at the designated spot. A photographic plate perpendicular to the optical axis and passing through the nominal focal point of the reflector intercepts the beam and produces a specified image. Eight such images are photographed on one plate at equally spaced intervals around the periphery of the zone being tested. A required number of zones are tested. Interpretation of the photographs is simple.

II. SCOPE

In order to determine the aberration characteristics of a convex surface, the problem that presents itself is how to determine the aberration characteristics when the incident light rays parallel to the optical axis of such a convex mold diverges after reflection. It is necessary to determine a method of causing these divergent rays to converge and, furthermore, to converge to a focus at a distance equal to the nominal focal length of the mold.

The theoretical solution of the problem is shown in Fig. 2. In this diagram, we see that a ray of light from the collimator* at position 1, parallel to the optical axis of the convex surface of a paraboloid of revolution, would, upon striking the mold, be reflected along a line which emanates from the virtual focus "f" of the mold. If a plane mirror "m_1," is placed with its reflecting plane in a vertical position somewhere along line HH' so that the reflected ray strikes "m_1" then the ray will again be reflected so that it passes through point "f'" if the following conditions are true:

a. That line HH' is tangent to the mold at a point where the optical axis pierces the surface of the mold.

b. That the reflected ray strikes mirror "m_1" on the line HH'.

If we examine triangle afm_1 and triangle $af'm_1$, we see that they are congruent since angle fm_1a equals angle $f'm_1a$ (law of reflection); right angle fam_1=right angle $f'am_1$; and am_1 is common to both triangles. This being true, $af=af'$; but af is the focal length of the mold; therefore, af' is also the focal length of the mold. By the same

* The collimator, so-called through long usage, is in reality a projector.

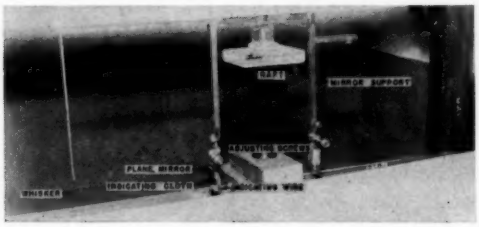

FIG. 3.

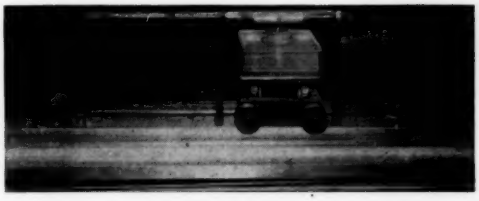

FIG. 4. Testing equipment for convex surface of paraboloid of revolution. Close up of assembly of raft and chassis on track showing how raft sits in container.

reasoning, if a ray suffers aberration (see path of ray r_i', position 2), it emanates from a virtual focal point ϕ which differs slightly from the established focal point and would be reflected through ϕ' which is exactly the same distance from "a" as ϕ. This shows that the amount of aberration is reflected exactly, so that a photographic plate "p" passed through f' would record the aberration of the convex surface of the mold in exactly the same manner as it would for a concave mirror.

There are several methods that may be applied to carry out the theoretical solution of the problem in a practical manner. However, whatever method is applied, it is necessary to make certain that:

a. Mirror "m" is in a vertical position at all times.

b. The degrees of freedom of motion of the mirror are limited only to that freedom of motion which permits the mirror to come freely to rest in a vertical position.

c. Mirror "m" is oriented in a plane which will permit the reflected image to intersect the optical axis of the convex mold under test. This is true when the plane is perpendicular to HH'.

d. Mirror "m" can be moved easily from one position to another along line HH'.

e. Mirror "m" is a front surface reflector to avoid errors due to refraction.

f. A cross hair mark be placed on the mirror that will permit the operator to see and place the image correctly in the dark.

g. The image on the mirror "m" will always lie in the plane tangent to the mold at the point where the optical axis pierces the surface of the mold.

III. DESCRIPTION OF EQUIPMENT

The equipment being used by the Engineer Board was built to conform to the aforementioned requirements as follows:

a. To keep the mirror in a vertical position, the following technique is used: A raft is floated in a pool of mercury. Since the mercury remains level regardless of how the container is tilted, the floating raft, in turn, will also be independent of the tilt of the containing vessel. If a mirror is fixed rigidly to the raft in a vertical position when the raft is afloat, the mirror will always be in a vertical position whenever it is again set afloat. In Fig. 3 we can see how the mirror supports are attached rigidly to the raft. It will be noticed that a mirror is attached to each end of the raft. The extra mirror in a plane parallel to the first one facilitates an important step in the examination of the mold. The important step (a step familiar to one determining the aberration of a 60-inch reflector) is that part of the test where a point on the mold is checked for identical aberration characteristics on opposite sides of the optical axis. This insures the accurate line-up of the apparatus with the optical axis of the mold. To do this with one mirror would mean turning the raft around 180° several times. To obviate this cumbersome step while the test is in progress two mirrors are used. While the raft is afloat in its container as shown in Fig. 4

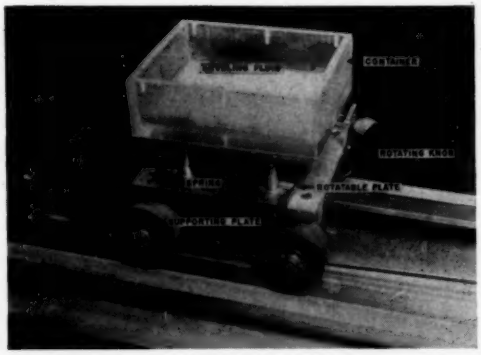

FIG. 5.

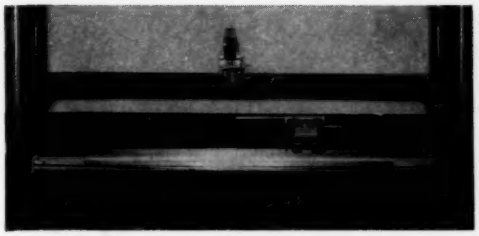

FIG. 6. Testing equipment for convex surface of paraboloid of revolution. Assembly of raft and chassis on track.

(see also Fig. 5 for component parts of container), the adjusting screws are set so that the mirrors are in a vertical position.

b. When the raft is floating in the container, its freedom of motion is limited only to the freedom to right itself to a horizontal and mirror "*m*" to a vertical position while it is floating. Any extraneous motion would nullify the accuracy of the test.

c. The reflected image is made to intersect the optical axis by manipulating the rotating knob shown in Fig. 5 so that the supporting plate moves the container and raft about a vertical axis which passes through the center of the plate. The image reflected from the mirror "*m*" intersects the optical axis when it is centered on a ground glass in the photographic plateholder.

d. To permit easy movement of the mirror from one position to another, the chassis is mounted on wheels which ride on a track, as shown in Fig. 6. The track is in a horizontal position.

e. To obtain a good first surface reflector, a small plane glass is aluminized.

f. The method used to make certain that the image is reflected from the correct position on the plane mirror while operating in the dark is based on the change of quality of light when reflected from a true reflecting surface and a diffuse reflecting surface. A strip of cloth is pasted on the mirror, as shown in Fig. 7(A). The operator, working in the dark, will see a sharp bright beam reflected from the mirror itself as shown in Fig. 7(A). If the chassis is moved on the track, the image starts to creep up or down and it will soon strike the cloth. The quality of light from the image is now changed from a bright one to a diffuse one. This will be quite apparent to the operator. The object is to place the image in a position where it just grazes the cloth as shown in Fig. 7(B). Should the image creep up only slightly on the cloth, as shown in Fig. 7(C), it will be detected because the quality of the light of the part of the image reflected from the cloth will differ from the bright portion of the image produced by the reflecting surface.

g. To make certain that the image on the mirror "*m*" will lie in a plane tangent to the mold at the point where the optical axis pierces

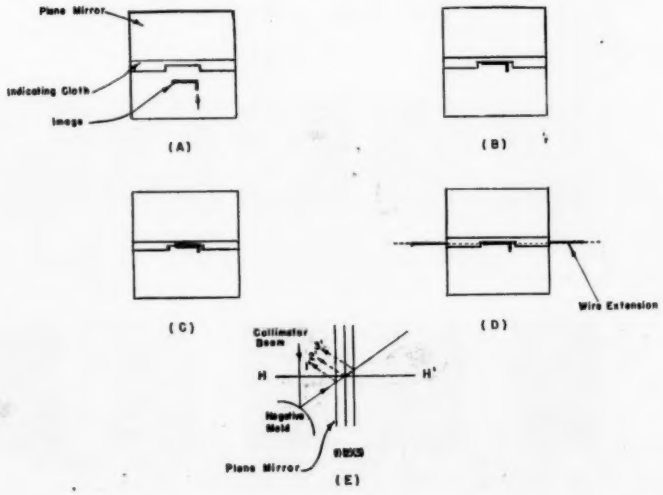

FIG. 7. Testing equipment for convex surface of paraboloid of revolution. Views showing placement of image on plane mirror.

FIG. 8. Apparatus for determining optical characteristics of a convex paraboloidal surface used as a mold in the electroforming of 60" searchlight reflectors. The collimator which projects a vertical beam of light onto the mold is not shown.

the surface of the mold, the following arrangement is used:

(1) Project wire extensions as shown in Fig. 7(D), on each side of the plane mirror in such a manner that an imaginary extension of the image in its correct position (adjacent to the indicating cloth) will coincide with the wires.

(2) To a supporting arm, shown in Fig. 3, a rod tipped with wire is attached in a vertical position. This wire can be raised or lowered in the chuck on the end of the rod. While the raft is afloat and the mirror set in a vertical position, the wire is set so that its point lies in the same horizontal plane as the wire extension in Fig. 7(D).

(3) With the track in a horizontal position, the wire is set directly over the crown of the convex mold as shown in Fig. 8 and lowered by the adjusting screw, until it just makes contact with the mold. At this point in the procedure, the crown of the mold and the point where the

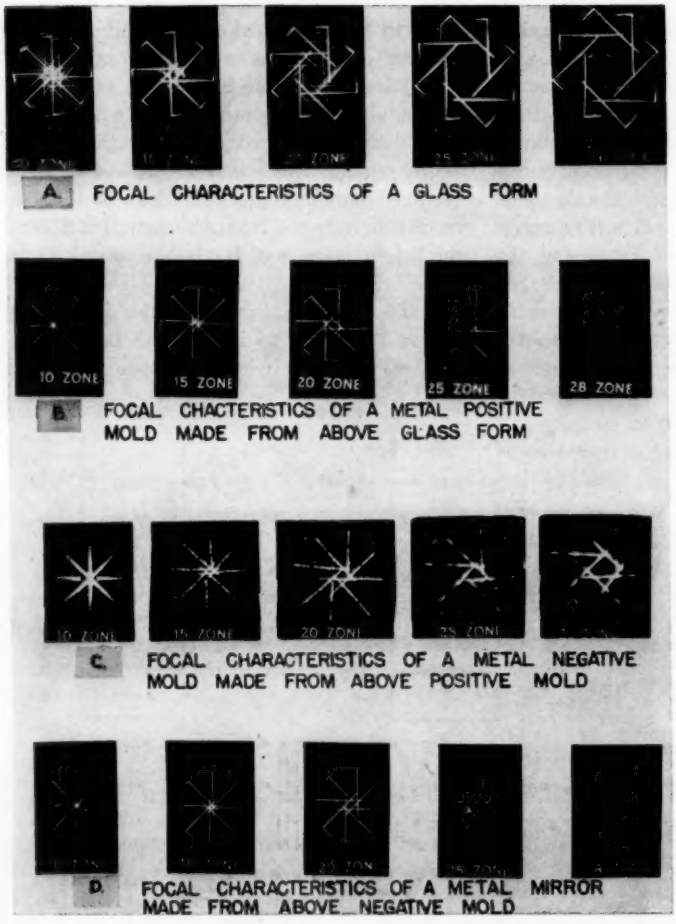
A. FOCAL CHARACTERISTICS OF A GLASS FORM
10 ZONE
15 ZONE
20 ZONE
25 ZONE
28 ZONE
B. FOCAL CHACTERISTICS OF A METAL POSITIVE MOLD MADE FROM ABOVE GLASS FORM
C. FOCAL CHARACTERISTICS OF A METAL NEGATIVE MOLD MADE FROM ABOVE POSITIVE MOLD
D. FOCAL CHARACTERISTICS OF A METAL MIRROR MADE FROM ABOVE NEGATIVE MOLD

FIG. 9. Focal characteristics.

image should fall on the plane mirror are in the same horizontal plane. We know now that when the operator moves the chassis so that the image falls, as shown in Fig. 7(D), the true picture of the aberration of the mold at that point is obtainable since the theoretical requirements are now met.

IV. PROCEDURE

If we again examine Fig. 8, we see that the convex mold rests on the turntable. The tracks are placed horizontally between the mold and the plateholder in such a manner that the end of the wire is close to the crown of the mold. The mold is then carefully centered and leveled in exactly the same manner as a concave mirror up to the point of preparation for the photographs of the aberration characteristics. From here on, the methods differ. In testing a concave mirror, the operator takes eight equally spaced pictures in each zone without any further qualifications, whereas in testing a convex mold, the operator must remember two important steps:

a. To set the wire directly over the crown of the mold and lower it by the adjusting screw, as shown in Fig. 8, until it just makes contact with the mold.

b. That the reflection of the image from the plane mirror must take place on a horizontal line tangent to the crown of the mold. Since the aberration in any particular zone is usually unsymmetrical, it is of *utmost importance* to place the plane mirror properly after each one of the eight pictures of the zone. This is to make certain that the reflection takes place from the proper position. Fig. 7(E) represents positions 1, 2, and 3 of the plane mirror normal to the line *HH'*. Reflected rays 1', 2', and 3' are possible from positions 1, 2, and 3 of the plane mirror. The only reflected ray that is correct is 2' which is reflected from the mirror in position 2, because that is where the reflection takes place on line *HH'*.

V. RESULTS

In Fig. 9, four aberration or focal tests are represented. They are the aberration characteristics of 60-inch parabolic glass form F 1751, metal concave mold No. 31, metal convex mold No. 66, and metal mirror No. 704, respectively. Metal concave mold No. 31 was produced from glass form F 1751. Metal convex mold No. 66 was produced from metal concave mold No. 31 and metal mirror No. 704 was produced from metal negative mold No. 66. There is a definite similarity in the focal tests of the concave mold, convex mold, and the mirror, but the focal test of the glass differs.

VI. DISCUSSION

The convex surface of the glass form F 1751 was used to make the concave mold. The aberration of this surface was measured by the standard apparatus *through the concave surface* of the glass to the silvered convex surface. However, the light beam suffers a refraction when it passes from the air into the glass and then out again after reflection from the back or silvered surface. This results in an apparently entirely different focal test from the molds and mirror it fathered. In order to obtain a correct picture of the focal characteristics of the glass form, it is necessary to perform the focal test directly from the convex surface.

VII. CONCLUSIONS

The above results prove that the determination of the aberration characteristics of a parabolodial convex surface is both possible and practicable. The advantages gained by such a test are:

a. The ability to determine directly the aberration of a convex mold.

b. The ability to check the change in aberration in a convex mold throughout its life.

c. The ability to establish a more precise aberration control in the production of mirrors.

A Study Concerning Characteristics of the High-Voltage a.c. Arc

E. A. BOETTNER* AND C. F. TUFTS**
Wyandotte Chemicals Corporation, Wyandotte, Michigan
(Received November 4, 1946)

This report covers the investigation of eight factors and their effects on the reproducibility of intensity ratios of spectral lines when the high voltage alternating-current arc is used as a source of excitation for quantitative spectrochemical analysis. The factors considered are photometry, weight of sample on the electrode, preburn time, exposure time, electrode separation, arc current, inductance, and anion-cation effect. Sodium hydroxide, sodium chloride, and calcium carbonate are employed for the work on the first seven factors. In the investigation of the anion-cation effect, eight inorganic salts are used. In conclusion there is a short discussion about data taken by a direct reading technique employing multiplier photo-tubes. Here an attempt is made to study short period variation in intensity ratio rather than the integrated effects produced when using photographic plates.

INTRODUCTION

THE high voltage a.c. arc was first suggested by Duffendack and Thompson,[1] while its application in the analysis of sodium hydroxide was described by Duffendack and Wolfe.[2] Oshry, Ballard, and Schrenk tell of its use in the analysis of dust, fumes, and ores.[3] This study reports certain effects that are produced when the arc is used as a source of excitation for quantitative spectrochemical analysis. Because of its sensitivity and reproducibility this source has found many applications, especially in the trace analysis of metals in chemical compounds. However, very little work has been done previously to determine the operating characteristics of this source and to determine reasons for the small fluctuations that are encountered when this source is used.

EQUIPMENT

The Arc Source

The a.c. arc, as installed in our laboratory, was designed to obtain maximum electrical stability. It has given trouble-free performance during the seven years that it has been in operation. It is possible that some of the inaccuracies reported in the literature when the a.c. arc is used, originate from inadequate electrical arrangements. For this reason it seems in order to include a description of the a.c. arc source as it is installed in this laboratory.

The circuit, which is of conventional design, is shown in Fig. 1. The transformer is a standard 15 kw distribution transformer operated at a maximum of 10 kw. The air-core inductance in the primary was wound with number 10 copper wire to produce sufficient inductance to maintain an arc across a 2-mm gap. The resistance was made from a continuous length of Hoskins number 20 Chromel "A" wire and was close-wound on a $\frac{1}{4}$ inch arbor. The coiled wire then was mounted on transite boards in the manner shown in Fig. 2, to permit free air passage about the wire. To provide the resistance desired, the wire was tapped at various places without making a break. The size of the wire was selected so that when operating at 5 amperes there was a mini-

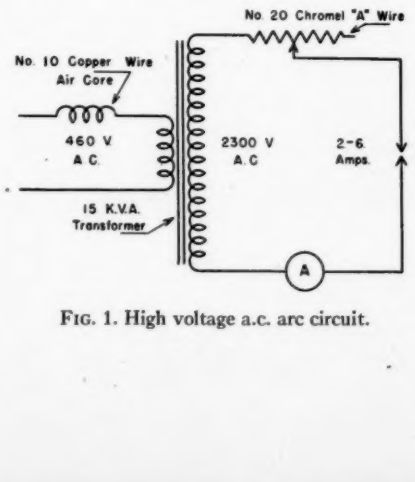

FIG. 1. High voltage a.c. arc circuit.

* Present address: Randall Laboratory of Physics, University of Michigan, Ann Arbor, Michigan.
** Present address: Armour Research Foundation, 35 West 33rd Street, Chicago 16, Illinois.
[1] O. S. Duffendack and K. B. Thompson, Proc. A.S.T.M. **36**, 301–309 (1936), Part II.
[2] O. S. Duffendack and R. A. Wolfe, Ind. Eng. Chem., Anal. Ed. **10**, 161–4 (1938).
[3] H. I. Oshry, J. W. Ballard, and H. H. Schrenk, J. Opt. Soc. Am. **31**, 627–33 (1941).

mum heating effect without employing an excessive amount of wire. Many installations use the wire-wound cone resistance because of availability and ease of installation. However, they have the disadvantage of operating at high temperatures and corroding at the point where the resistance wire joins the outlet terminals. The source is capable of operating up to 5 amperes at 2300 volts or 2.2 amperes at 4600 volts.

The arc stand is similar to that used by Duffendack and Wolfe.[2]

Other Equipment

Unless otherwise stated, the spectrograph, auxiliary equipment, photometric method, and sample preparation are the same as described previously.[4]

PRECISION IN THE ANALYSIS OF SODIUM HYDROXIDE

The alternating-current source in our laboratory has been used daily for seven years for the determination of metallic impurities in alkali products. During this time more than 22,000 samples have been analyzed and more than 80,000 determinations made. The precision of the results has been consistent during this period and is still in close agreement with the values published previously.[4]

Sodium hydroxide was used for the greater portion of this study because our laboratory has accumulated a large amount of precision data about this matrix, and because the best precision attained has been in the analysis of sodium hydroxide.

FIG. 2. Arc resistance bank.

[4] E. A. Boettner, Ind. Eng. Chem., Anal. Ed. **13**, 861–4 (1941).

TABLE I. Precision and range of analyses.

	Spectral lines Impurity A	Mo. A	Range of analysis %	Average deviation from mean %
		25% sodium hydroxide		
Si	2881.6	3158.2	0.0005 –0.0050	3.2
Si	2987.7	3158.2	0.0190 –0.0700	4.2
Al	3092.8	3158.2	0.00015–0.0010	2.0
Al	3082.2	3158.2	0.0002 –0.0015	2.6
Ca	3968.5	3903.0	0.0001 –0.0015	4.5
Mg	2802.7	2816.2	0.0004 –0.0010	4.3
Fe	2483.3	2649.5	0.00007–0.0005	5.0
Mn	2794.8	2649.5	0.00002–0.00005	3.0
Cu	3247.6	*a*	0.00004–0.0002	2.5
Ni	3002.5	*a*	0.00008–0.0006	2.6
Ni	3050.8	*a*	0.00008–0.0006	3.2
Sr	4077.7	3903.0	0.0001 –0.0020	7.8
Nine elements				3.7

a = Background next to line.

Table I shows the spectral lines used, the ranges of analyses, and the precision of the analyses expressed as the average percent deviation from the mean. The average percent deviation from the mean for nine elements is 3.7 percent. This value represents the average deviation from the mean of results for 2 or 3 spectrograms run on the same plate. Because the data originate from individual plates, errors due to plate processing and plate calibration do not enter into these figures or into any of the data presented in this paper. Errors in microphotometry and calculation, however, are included in the deviations, and are discussed next.

PHOTOMETRY

In considering photometry as a factor, the concern is about deviations acquired in plate microphotometry and in calculating technique. It was necessary to establish the fact that these two operations did not contribute significant errors to the precision values in Table I. To test our precision in this connection a plate containing 24 spectrograms of sodium hydroxide was employed. Three observers performed the experiment by measuring the 2987.7A silicon line (impurity) and the 3158.2A molybdenum line (internal standard) in each spectrogram. Each observer determined the blackening of the lines twice and converted both sets of values

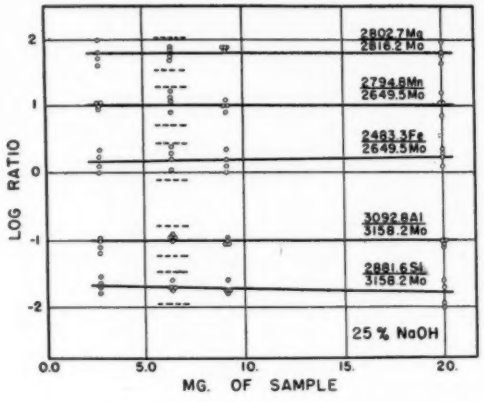

FIG. 3. Effect of sample weight on log ratio.

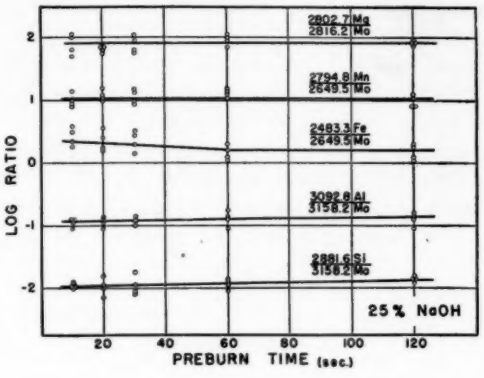

FIG. 4. Effect of preburn on log ratio.

into percent of silicon. All readings were made on the same day with at least one hour elapsing between the readings of each observer. One pattern curve was used for all calculations to convert blackening values to percent silicon. The difference between the observer's two values of percent silicon for any spectrogram was divided by the mean for these two values to determine the percent difference. There were 24 such values for each observer and the average of these, expressed in average percent difference, was ± 1.25 percent.

This relegates the origin of the greater portion of the 3.7 percent deviation to one or both of the following:

(a) The preparation of electrodes;
(b) Conditions that arise when the sample is burned in the arc.

In the studies that follow, only one factor was varied at a time, holding all others constant, and at the value used in making control analyses. The standard conditions used for analyzing sodium hydroxide are as follows:

Preburn:	30 seconds
	½ mm gap
	2.7 amperes
	2300 volts
	1.7 millihenrys
Exposure:	60 seconds
	1 mm gap
	4.8 amperes
	2300 volts
	1.7 millihenrys

THE WEIGHT OF SAMPLE ON THE ELECTRODE

In the preparation of the sample, the weight of material coated on the electrode is the only important factor to be considered. Treatment of the carbon electrode with kerosene or a similar sealing agent did not affect precision. The viscosity of 25 percent sodium hydroxide is such that drying is complete before penetration of sample into the electrode can exert a noticeable effect. Figure 3 shows how variation in sample size affects the log ratios of five impurity lines in 25 percent sodium hydroxide. The short dotted lines in the figure represent the variation in log ratio necessary to produce a deviation of ± 10 percent in the analysis. It is evident that this factor is not affecting the precision.

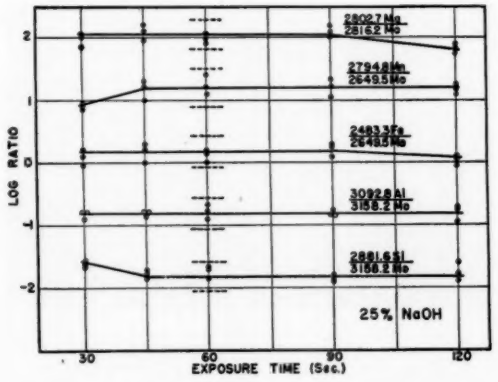

FIG. 5. Effect of burning time on log ratio.

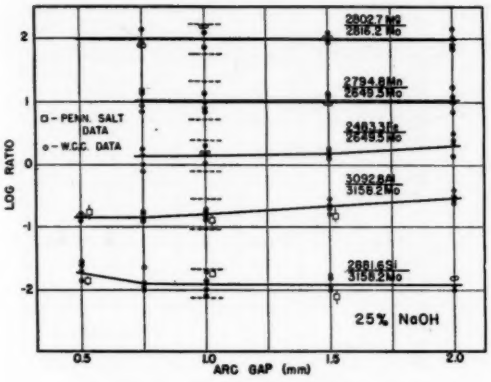

FIG. 6. Effect of arc gap on log ratio.

PREBURN TIME

Figure 4 considers the effect produced when the time of preburning is made variable. Again the variable produces a negligible effect.

EXPOSURE TIME

Exposure time, when varied, caused no significant effect. Figure 5 shows the result for this part of the investigation.

ELECTRODE SEPARATION

Figure 6 shows the effect of electrode separation on log ratio. As can be seen from the figure, changes in the gap have little effect on the ratio, which is somewhat contrary to what has been published previously. Because of this it was decided that this factor should be checked in another laboratory equipped similarly. Mr. Karl W. Beyer, of the Pennsylvania Salt Manufacturing Company, kindly agreed to repeat this experiment. The Wyandotte Laboratory of the Pennsylvania Salt Manufacturing Company is equipped with a medium quartz spectrograph and a source similar to that described previously in this paper. The data Mr. Beyer furnished have been incorporated in Fig. 6 and although it shows some difference in the case of the silicon line, in general, there is good agreement between the two laboratories.

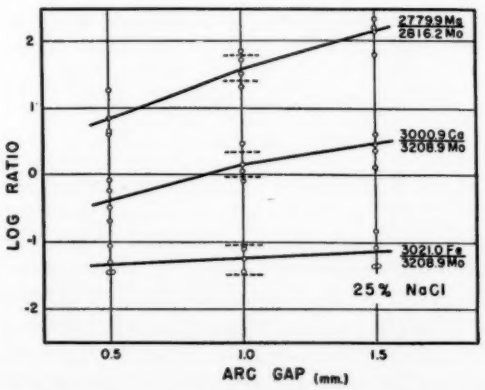

FIG. 7. Effect of arc gap on log ratio for sodium chloride.

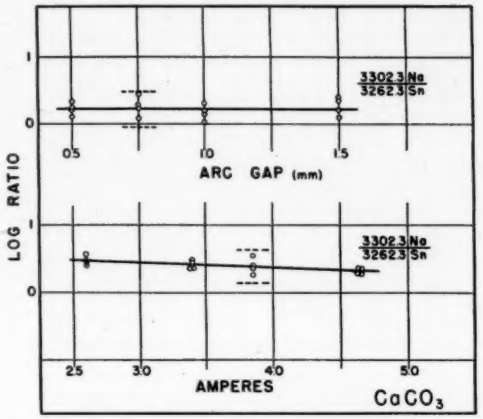

FIG. 8. Effect of gap and amperage on log ratio for calcium carbonate.

The work on electrode separation was continued in our laboratory by employing a sample of 20 percent aqueous sodium chloride solution. Figure 7 shows the results of this experiment. Here the deviation of log ratios for magnesium and calcium was more pronounced. To determine whether this variation was sufficient to affect our analysis, a short study was made to establish the accuracy with which the arc gap could be set. By employing a projected image of the electrodes and gap so as to produce an enlargement of approximately 45 times, it was found that the average deviation from an arbitrary standard gap was ± 0.05 mm with a maximum deviation of 0.11 mm. To complete this arc-setting study, a plate containing 24 spectrograms of sodium chloride was exposed with the arc gap of each exposure measured by using the projection device. The results obtained showed no correlation between the variation in log ratio and the small variations in arc gap.

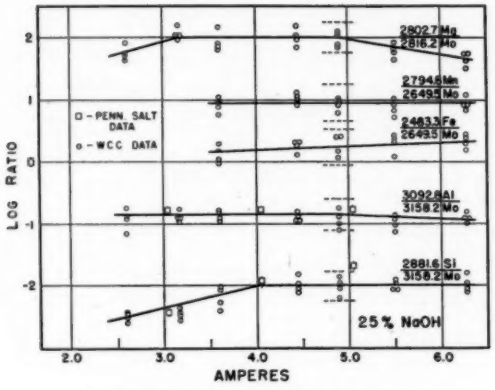

FIG. 9. Effect of amperage on log ratio.

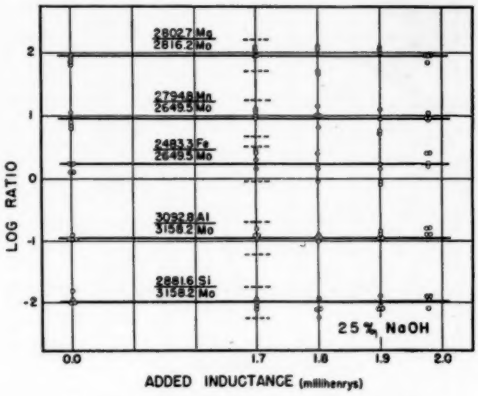

FIG. 10. Effect of inductance on log ratio.

Figure 8 (upper curve) shows the electrode separation effect using calcium carbonate. This is a special case as far as sample preparation is concerned since the solid calcium carbonate is dispersed in a 10 percent nitric acid solution to form a slurry. This chemical was tested because it shows excellent reproducibility in routine spectrochemical analysis for sodium. Here arc gap has no effect upon reproducibility for the range considered.

When comparing the data given here for sodium hydroxide with that reported in the original paper of Duffendack and Wolfe[2] it should be remembered that they were limited in their work to an arc current of 2.5 amperes without added inductance. It is apparent when comparing the two studies that the effect of arc gap on log ratio is much greater when using amperages less than 3.0.

ARC CURRENT

The lower curve in Fig. 8 shows the effect of arc current variations for calcium carbonate. Figure 9 shows the effect of varying arc current using 25 percent sodium hydroxide as the sample material. Data from the Pennsylvania Salt Laboratory are included also in this study. Both laboratories show good agreement with data published previously by Duffendack and Wolfe.[2]

The data in the two figures show that small variations (less than 0.5 ampere) in the arc current do not affect the log ratio appreciably. In this laboratory the amperage is constant to 0.1 ampere.

INDUCTANCE

Figure 10 shows inductance as the variable. Here again, is a factor that contributes practically no effect to log ratio variation when using the high voltage a.c. arc.

ANION-CATION EFFECT

In Fig. 7 where the data deal with sodium chloride it is quite noticeable that the spread of points is greater than in figures dealing with sodium hydroxide. The average deviation for impurities in sodium chloride is about twice as great as for sodium hydroxide. This condition suggested the possibility that certain anions and cations might affect reproducibility. Table II lists the eight salts used in an anion-cation study. Values for average deviation represent data from 354 determinations using the following log ratios:

$$\frac{3082.2\ \text{Al}}{3158.2\ \text{Mo}} \quad \frac{3092.8\ \text{Al}}{3158.2\ \text{Mo}} \quad \frac{3968.5\ \text{Ca}}{3903.0\ \text{Mo}}$$

$$\frac{2802.7\ \text{Mg}}{2816.2\ \text{Mo}} \quad \frac{2881.6\ \text{Si}}{3158.2\ \text{Mo}}.$$

These ratios have been tested repeatedly in our sodium hydroxide analysis and have been found to be relatively good. The lines measured originated from impurities in the salts. The impurities present in all cases were less than 0.1 percent. Results seem to indicate a lack of definite trends that favor one cation, one anion, or one compound. The only exception was copper chloride where the reproducibility was consistently poor.

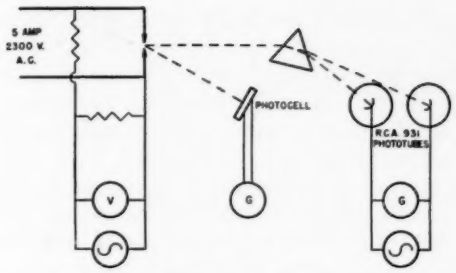

FIG. 11. Instrumentation for study of short-period variations.

FIG. 12. Equipment for observing short-period variations.

The relatively small amount of data employed in this section of the study makes it necessary to reserve final judgment until more determinations can be considered.

SHORT-PERIOD VARIATIONS

It is apparent from the data accumulated on photographic plates that the fluctuations or variations in question have a time-period too short to detect photographically. For this reason it was decided to use photo-tubes as spectral line receivers.

A schematic diagram of the equipment used is shown in Fig. 11. Two electron multiplier photo-tubes were attached to the spectrograph to receive the radiation of two lines in the manner described in a previous paper.[5] Across the opposed output of the two tubes was placed both a galvanometer and a cathode-ray oscilloscope. The lines used were the 2882A silicon and 3961A aluminum in a sodium hydroxide matrix.

Another photo-cell, of the barrier layer type, was exposed to receive the total radiation of the arc. It was connected to a microammeter, showing variations in the light output.

A voltmeter and another oscilloscope were placed across the arc gap to show variations in voltage drop across the gap.

A picture of the entire unit is shown in Fig. 12. The meters and oscilloscope screens were grouped to include them all in the field of a single motion picture camera. It then was possible to record their action photographically during the burning of the arc, and to examine the resulting film later.

It was hoped that by obtaining data in this manner, it would be possible to trace any variations in the intensity ratio that might originate either from the fluctuating voltage drop across the arc or from changes in light intensity of the arc.

A study of several hundred feet of motion picture film failed to correlate the variations to these sources, but definite conclusions cannot be drawn without making some changes in the set-up described. It has been found that the use of meters such as the voltmeter and microammeter used here is not desirable because of their relatively long time period as compared to the rapid fluctuations being studied. Changes in the circuit using the photo-multiplier tubes are also desirable.

TABLE II. Precision of analysis of eight inorganic salts.

Compounds tested	% average deviation
NaCl	7.7
$NaNO_3$	10.0
NH_4Cl	13.0
NH_4NO_3	8.7
$CuCl_2$	24.0
$Cu(NO_3)_2$	9.0
$MgCl_2$	7.3
$Mg(NO_3)_2$	12.0

The above table was compiled from 354 determinations by using the following log ratios:

$$\frac{3082\ \text{Al}}{3158\ \text{Mo}};\ \frac{3093\ \text{Al}}{3158\ \text{Mo}};\ \frac{2803\ \text{Mg}}{2816\ \text{Mo}};\ \frac{2882\ \text{Si}}{3158\ \text{Mo}};\ \frac{3969\ \text{Ca}}{3903\ \text{Mo}}.$$

FUTURE WORK

From the data presented to this point, it can be seen that the variations in question are not traceable to the factors investigated in this report. The work is being continued to study

[5] E. A. Boettner and G. P. Brewington, J. Opt. Soc. Am. **34**, 6–11 (1944).

further its characteristics. The anion-cation effect, especially of the alkali and alkali earth elements will be investigated, as will also the background produced by some of these chemicals. Work has already been done on potassium hydroxide, which acts quite differently from sodium hydroxide in the a.c. arc and much has been learned concerning the reduction of the intense background produced by this alkali. The reduction of background was affected in this case by the addition of 10 percent alcohol. The effect of organic solvents on inorganic salts will be studied further.

It is hoped that time will permit the modification of the photo-tube recording to study further the short-period variations in the arc.

CONCLUSION

1. Variations in weight of sample on electrode, preburn time, exposure time, and added inductance have no effect on log ratio in the analysis of sodium hydroxide.

2. Small variations in arc gap have no effect on log ratio in the analysis of sodium hydroxide, sodium chloride, or calcium carbonate. However, a gap of 1.0 ± 0.1 mm is used to obtain spectra of uniform intensity on the photographic plate.

3. Small variations in current through the arc have no effect on log ratio in the analysis of sodium hydroxide and calcium carbonate. However, the current is held to within ± 0.1 ampere to obtain spectra of uniform intensity on the photographic plate.

4. In testing eight salts to determine the effect of anions or cations on the precision of analysis, all were found to be similar with the exception of copper chloride, whose reproducibility is very poor.

5. Modifications in the apparatus used to make the short period study are necessary before definite conclusions concerning this portion of the study can be reached.

ACKNOWLEDGMENT

The authors thank Dr. G. P. Brewington, for advice in setting up the apparatus for the short-period variations study; Dr. R. A. Wolfe for suggestions throughout the investigation; Mr. Karl W. Beyer and the Pennsylvania Salt Manufacturing Company for contributing data on electrode separation and arc current; and Mr. B. D. Field, who assisted in the initial phase of the study.

Conference on Technique of Measurements in Stockholm

THE Royal Swedish Academy of Engineering Sciences, and the Association of Technical Physicists are planning a conference on instruments and measurements to be held in Stockholm June 4–7, 1947. The program will contain symposia on Industrial Spectroscopy, Testing of Materials, Industrial Control, and Metrology. The Swedish Society for Metallography will cooperate in the symposia on the first two subjects. Membership in the conference requires a subscription of 10 Swedish crowns (about $2.80). Those interested in attending or in obtaining further information should write Mr. Owe Berg, Mariedalsvägen 27, Huddinge, Sweden.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 3 MARCH, 1947

Optimum Spectral Sensitivities for a Reversible Color Film Process

RONALD H. BINGHAM AND HERMAN HOERLIN
Research Laboratories, Ansco, Binghamton, New York

(Received October 14, 1946)

A method was developed for theoretical evaluation of spectral sensitivity functions to be used with a specified set of subtractive primary dyes in a reversible color process. Using a numerical calculus of variations procedure, hypothetical blue, green, and red sensitivities were varied in a systematic way to obtain the curves which minimize the errors of color reproduction. These three optimum sensitivities were found to be non-overlapping and to have maxima at 450, 540, and 630 mμ, respectively. As an illustration of the application of the method to the evaluation of experimental sensitizers, color renditions are calculated for two sets of sensitivity functions which are alike in the blue and green, but different in the red. The two combinations of sensitivities are compared and conclusions are drawn regarding changes of the curves which would yield specific improvements in color reproduction.

I. INTRODUCTION

ACCORDING to the colorimetric theory of additive three-color reproduction as outlined by Hardy and Wurzburg,[1] each of the spectral sensitivity curves of the three photographic emulsions used in recording the original scene should be a linear transformation of the tristimulus value functions of normal human vision. The relationship between the trichromatic specifications of the projection primaries and the required sensitivity curves is such that for any real set of primaries, the spectral sensitivities are in each case, negative over parts of the spectrum. MacAdam[2] examined several compromises according to which negative spectral sensitivity values could be avoided, and concluded that, in an additive process, the simple expedient of using only the positive parts of the theoretically required curves and neglecting the negative portions would lead to the smallest errors.

MacAdam also analyzed the subtractive process,[3] developing a method for associating a given set of three subtractive primary dyes with a corresponding set of three hypothetical additive primaries. This procedure yielded the specifications of an additive process which was closely equivalent to the subtractive process in question, and for which the required spectral sensitivities could be determined. The analysis led to a set of spectral sensitivity functions which were partly negative and which were characterized by a considerable amount of overlap. The need for masking was established as a means of producing the effect of negative sensitivities and in addition, to correct for the overlap of the subtractive dyes. In applying these results to the wash-off relief process, it was found that the negative parts of the spectral sensitivity curves were relatively small and could be neglected without incurring serious errors.

The general procedure in the above investigations was to derive the spectral sensitivity functions from the standard data of additive colorimetry by assuming perfect color rendition. In modern subtractive color processes of the reversible type, however, masking is difficult. For this reason, it becomes important to investigate the conditions for optimum reproduction under restrictions which preclude the possibility of perfection. The problem was, therefore, phrased as follows:

> *Using a given set of subtractive primary dyes, what set of spectral sensitivity functions would yield the closest approach to perfect reproduction, assuming that negative sensitivities and masking are to be excluded?*

The problem, thus stated, is essentially one of calculus of variations, in which three empirical functions are to be simultaneously varied in order to minimize the average differences between all possible original and reproduction colors. The empirical form of the functions involved as well

[1] A. C. Hardy and F. L. Wurzburg, Jr., "Theory of three-color reproduction," J. Opt. Soc. Am. **27**, 227 (1937).
[2] D. L. MacAdam, "Photographic aspects of the theory of three-color reproduction," J. Opt. Soc. Am. **28**, 399 (1938).
[3] D. L. MacAdam, "Subtractive color mixture and color reproduction," J. Opt. Soc. Am. **28**, 466 (1938).

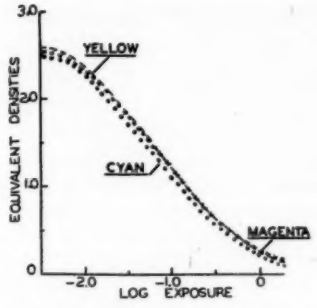

FIG. 1. Characteristic curves of typical, nearly balanced, reversible color film $\gamma = 1.2$.

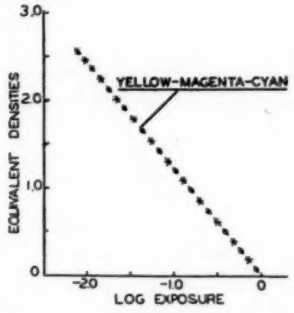

FIG. 2. Characteristic curves of idealized, perfectly balanced, reversible color film at $\gamma = 1.2$.

as the complexity of the color difference formulas when expressed as functions of the spectral sensitivities made it necessary to resort to numerical integration methods as a means of solution.

A number of representative colors were specified by selected wave-length charts in a form which facilitated the calculation of the corresponding color renditions with any desired set of spectral sensitivity functions. In order to make a systematic study of various sets of sensitivities, each sensitivity function was formulated as the sum of a series of "bell-shaped" curves in much the same way that a periodic function may be expressed as a sum of sine and cosine terms. By varying the relative contributions of the several bell-shaped curves, practically any desired sensitivity function could be obtained.

Reproductions were calculated for each of eleven color samples for a large number of spectral sensitivity functions. The set of sensitivities which was selected as optimum was that for which the over-all color distortion was found to be a minimum.

Before considering in detail the procedure which was used, it is necessary to discuss certain concepts of color sensitometry and to state the basic theory of subtractive color reproduction as it bears upon the problem.

II. COLOR SENSITOMETRY

The concept of equivalent density was defined by Heymer and Sundhoff[4] and by Evans[5] essentially as follows: In a subtractive color process, the equivalent density of one of the primary dyes of a reproduction is equal to the minimum gray density obtainable by superimposition upon those amounts of the other two primaries which are necessary in order to yield visual gray. In the present paper the gray density involved in this definition will be defined as follows:

$$d = \log_{10} F_i / F_t,$$

where F_i is the luminous flux normally incident upon the sample and F_t is the luminous flux transmitted. We will designate the equivalent densities in the yellow, magenta and cyan layers by j, m, and c respectively.* Methods of determining the equivalent densities of individual layers from measurements on superimposed three layer color film have been discussed in the literature.[5,6]

The three sensitometric curves of a typical, nearly balanced, reversible film are shown in Fig. 1 while in Fig. 2 are shown the corresponding curves of the idealized, perfectly balanced material which forms the basis of the present calculations. A contrast value, $\gamma = 1.2$ is used throughout the present work, in order to approximate the intensity relationships which characterize the projected images yielded by typical commercial color processes. It is convenient, following MacAdam,[2] to measure the exposure in inertia units; that is, the value of log exposure for each

[4] G. Heymer and H. Sundhoff, "Über die Messung der Gradation von Farbenfilmen," Agfa Zentral-Labor. Veröff. 5, 62 (1937).

[5] R. M. Evans, "Color densitometer for subtractive processes," J. Soc. Mot. Pict. Eng. 31, 194 (1938).

* In discussions of subtractive color reproduction, we use the letter j instead of y to denote yellow equivalent density in order to avoid conflict with the notation of additive colorimetry.

[6] R. H. Bingham, "Sensitometric evaluation of reversible color film," J. Soc. Mot. Pict. Eng. 46, 368 (1946).

of the three layers is taken as zero at the point where the straight-line portion of the sensitometric curve intersects the horizontal axis. The introduction of such a definition constitutes the mathematical equivalent of balancing the emulsion.

The spectral density curves for the three layers of a sample of unit equivalent density are experimentally determined functions of wave-length which will be designated by J_1, M_1, and C_1. The curves which were used in these calculations are shown in Fig. 3 and are representative of the subtractive dyes which were in use in a reversible color process prior to 1941. We now assume that, for the primary dyes within the equivalent density range 0–2.5, the spectral densities are given by jJ_1, mM_1, and cC_1, respectively.* Thus if the spectral density of the three-layer sample is represented by D, we have,

$$D = jJ_1 + mM_1 + cC_1. \tag{1}$$

The corresponding spectral transmittance is given by

$$T = 10^{-D} = \log^{-1}(-jJ_1 - mM_1 - cC_1). \tag{2}$$

The concept of spectral sensitivity is discussed by MacAdam.[2] We shall assume, as he does, that the contrasts of the individual layers are independent of wave-length and shall define the spectral sensitivity of a layer as the reciprocal of the amount of energy at the specified wave-length necessary to yield zero equivalent density. We thus have three spectral sensitivity curves which characterize the yellow, magenta, and cyan layers and which may be referred to as the blue, green, and red sensitivities, respectively.

III. THEORETICAL DEVELOPMENT

Fundamental colorimetric and photographic theory provides the basis for a formal solution of the problem even though direct application would necessitate extremely laborious calculations. We may first present those basic equations of color reproduction, which are needed for the determination of optimum sensitivities, leaving the details of computation until later.

Let us assume that we wish to photograph a sample of spectral reflectance R, which is irradiated by a source of spectral energy U, both R and U thus being functions of wave-length. We may specify the color in the usual way* by the tristimulus values X', Y', and Z':

$$X' = \int_0^\infty URx\,d\lambda,$$

$$Y' = \int_0^\infty URy\,d\lambda, \tag{3}$$

$$Z' = \int_0^\infty URz\,d\lambda.$$

In these equations, x, y, and z are the tristimulus values of the spectrum colors as defined for the I.C.I. standard observer. Using van Kreveld's law of summation of photographic exposures[7,8] our above definition of the spectral sensitivities yields the following expressions for the blue, green, and red exposures:

$$E_b = \int_0^\infty URS_b\,d\lambda,$$

$$E_g = \int_0^\infty URS_g\,d\lambda, \tag{4}$$

$$E_r = \int_0^\infty URS_r\,d\lambda,$$

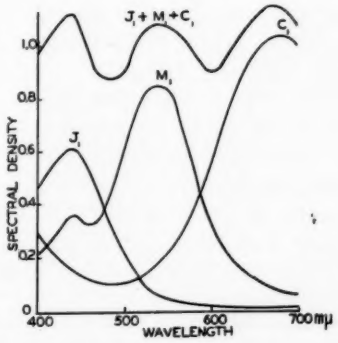

FIG. 3. Spectral density curves of primary dyes at unit equivalent densities.

* This amounts to assuming a restricted form of Beer's law which holds satisfactorily well for the dyes of the type under discussion.

* Tristimulus values and equivalent densities which are characteristic of the sample are primed to distinguish from those of the reproduction.

[7] A. van Kreveld, "Ein Gesetz für die Schwärzungskurven bei Mischfarben," Zeits. f. wiss. Phot. **32**, 222 (1934).

[8] J. H. Webb, "Summation of different color radiations by a photographic emulsion," J. Opt. Soc. Am. **26**, 12 (1936).

where S_b, S_g, and S_r are the spectral sensitivities. For the straight-line portion of the sensitometric curves, the equivalent densities of a reproduction are given by

$$j = -\gamma \log E_b,$$
$$m = -\gamma \log E_g, \quad (5)$$
$$c = -\gamma \log E_r,$$

where γ is the contrast of the material.** If such a reproduction is viewed by the same illuminant as was used during exposure, its color is specified by

$$X = \int_0^\infty U\bar{x} \log^{-1}(-jJ_1 - mM_1 - cC_1)d\lambda,$$
$$Y = \int_0^\infty U\bar{y} \log^{-1}(-jJ_1 - mM_1 - cC_1)d\lambda, \quad (6)$$
$$Z = \int_0^\infty U\bar{z} \log^{-1}(-jJ_1 - mM_1 - cC_1)d\lambda.$$

Equations (4)–(6), together give the tristimulus values X, Y, and Z of the reproduction in terms of the spectral characteristics of the three subtractive primary dyes, the spectral sensitivities of the three photographic layers, the spectral reflectance of the sample, and the spectral energy distribution of the light source.

Had our problem been merely to determine the color distortion which would result from the use of a specific set of three spectral sensitivity functions, this could have been done by selecting a set of representative color samples and determining in each case the tristimulus values of sample and reproduction. The graphical representation of these data on the I.C.I. chromaticity diagram would have given indications of the difference between originals and reproductions. As soon as we consider analyzing a large number of different blue, green, and red sensitivity curves in all possible combinations, however, this straightforward procedure becomes hopelessly laborious so that a different approach is necessary.

IV. SUBTRACTIVE COLOR SPECIFICATION BY USE OF EQUIVALENT DENSITIES

In the present problem, we were faced with the necessity of analyzing a large number of sensitivity curves in different combinations. Were we to carry out the analysis in the manner described above, it would necessitate numerical integration of Eqs. (6) for various combinations of j, m, and c given as functions of S_b, S_g, and S_r. Such a procedure would leave us at the disadvantage of not knowing whether or not a particular sensitivity function was near optimum until it had been evaluated in triple combination with a considerable number of other sensitivities. The number of integrations might thus be expected to increase nearly as the third power of the number of sensitivities to be tested.

As it happened, however, a method of analysis was developed which obviated this difficulty. Referring to Eqs. (4) and (5) we see that a choice of a particular spectral sensitivity curve uniquely determines the corresponding equivalent density, assuming that the conditions of exposure and processing remain constant. We may also note that, for any combination of j, m, and c, the tristimulus values X, Y, and Z of the reproduction are determined as long as the spectral density functions of the primary dyes remain fixed. These observations suggest that the variables j, m, and c are, for a particular color process, valid for purposes of specifying the color of a reproduction. As we shall see, their use in this manner greatly reduces the number of combinations of sensitivities which have to be investigated.

In order to carry out the analysis in terms of equivalent densities, however, it was necessary to have a corresponding specification for each original sample for which color rendition was to be determined. This was done by determining, for each color sample, its "effective" equivalent densities j', m', and c', i.e., those equivalent densities of a reproduction which would just perfectly match it.

Then in analyzing the series of blue sensitivity combinations, for example, values of the reproduction equivalent densities j were computed for each sensitivity combination for each of the color samples. A comparison of these values with the corresponding value of j' for the samples led to the elimination of all but the few most promising sensitivities which were near optimum. Finally, comparing all data on near-optimum sensitivities, the set which gave the best over-all

** The right-hand members of Eqs. (5) are negative indicating that density increases with decreasing exposure.

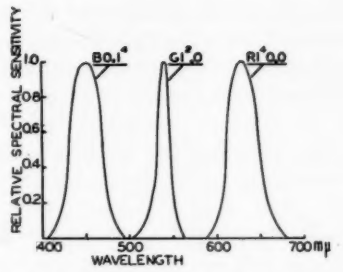

FIG. 4. Typical blue, green, and red component spectral sensitivity curves with maxima at 450, 540, and 630 mμ.

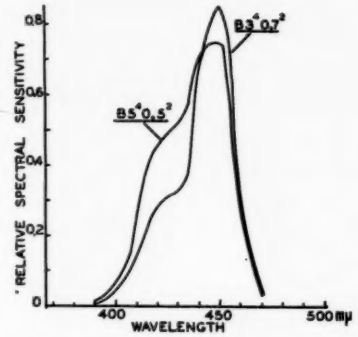

FIG. 5. Blue spectral sensitivity curves formed by combination of two components.

reproduction was selected as optimum for use with the three primary dyes for which the calculations were made.

V. FORMULATION OF SPECTRAL SENSITIVITY CURVES

The above considerations indicate that the mathematical solution of the problem would require the specification of spectral sensitivity functions which could be varied in a systematic way. The form of specification had to be such as to permit convenient and rapid integration of Eqs. (4) according to which the photographic exposures were computed.

As a means of formulating sensitivity curves which could be varied systematically, a series of bell-shaped component sensitivities was defined from which curves of practically any shape could be synthesized. One of the curves shown in Fig. 4 represents a typical blue component sensitivity. All members of the blue series were similar in shape but differed in position of the maximum as well as in width at half maximum. Maxima were taken at ten millimicron intervals, while widths were taken as either ten, twenty, forty, or sixty millimicrons. In formulating a sensitivity curve, simple decimal fractions of two or more components were taken; for example, a blue sensitivity might consist in six parts of the component at 430 mμ of width 40 millimicrons and four parts of the component at 450 mμ of width 20 millimicrons. Such a curve is designated $B6^{4}0.4^{2}$ with the letter distinguishing among blue, green, and red sensitivities. Each digit in this designation represents the relative height of a sensitivity component, while the exponent indicates its width. The spectral position of each component is indicated by the location of the digit relative to the decimal point which separates the 440 mμ and 450 mμ positions.

A curve composed of a single component is denoted in the same way with "1" being used as the digit. Thus, if the two components referred to above constituted two separate sensitivity curves, they would be designated as $B1^{4}0.0$ and $B0.1^{2}$, respectively.

Also shown in Fig. 4 are typical sensitivity components in terms of which the green and red sensitivity curves were defined. The same system of designation was followed with the letters G and R indicating the colors. In the case of a green sensitivity curve, the decimal point in the designation separates the 540 mμ and 550 mμ position, while, in a similar way the decimal point in the designation of a red sensitivity curve separates the 640 mμ and 650 mμ positions.

Combination blue and combination red sensitivities were formulated and studied. Their curves are shown in Figs. 5 and 6. It was found in the case of the green, however, that single component sensitivities were very nearly optimum so it was unnecessary to formulate combination curves.

VI. DEFINITION OF COLOR SAMPLES

As indicated in Section IV, the method of analysis required that each color sample for which renditions were to be determined, be specified in terms of its effective equivalent densities—i.e., in terms of the values of j', m', and c' which would characterize perfect color reproduction. In addition, we should note that, in

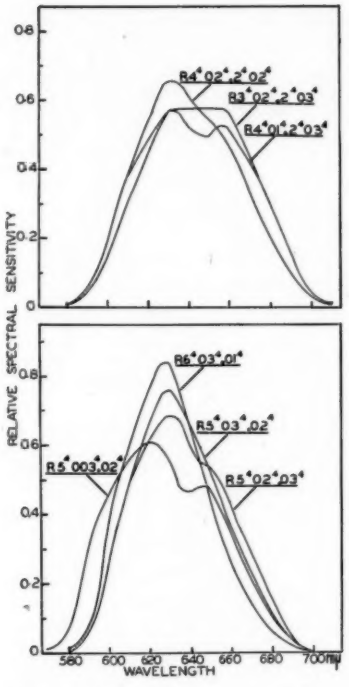

FIG. 6. Red spectral sensitivity curves formed by combination of three or more components.

order to be able to determine the reproductions of these color samples, their spectral reflectance curves had to be specified in a manner which would permit the convenient integration of Eqs. (4).

In defining the set of color samples which was used, it was considered important that several colorimetrically homologous pairs be included. Thus, two blue samples of identical color but of different spectral reflectances—one reflecting strongly in the far red, for example,—should be expected to be reproduced most nearly the same by optimum sensitivities. This principle constituted one of the criteria for deciding among the many sensitivities which were studied.

Before discussing the method of sample specification which was employed, let us consider, briefly, two other possible procedures. While each of these methods appear at first glance to be more straightforward than the one which was actually used, each involves more laborious computation and makes the selection of homologous pairs either impossible or very difficult.

One possible method would have been to define each color sample arbitrarily in terms of its effective equivalent densities and then consider its reflectance to be numerically equal to the transmittance of the corresponding perfect reproduction. This amounts to formulating the reflection samples from combinations of the dyes of the process and precludes the possibility of including homologous pairs, since three dyes can be combined in only one combination to yield any arbitrarily specified color. In addition, the spectral characteristics of different samples defined by such a procedure would have maxima and minima at approximately the same wave-lengths and would therefore be less typical of normal reflecting samples.

A second method which was considered but rejected was to specify the reflectance curves of a number of samples and then to calculate their corresponding effective equivalent densities. To find homologous pairs by this method would be possible but would be a laborious trial-and-error process. In order to calculate the effective equivalent densities of such a sample it would have been necessary to integrate Eqs. (6), several times, trying different sets of values of j', m', and c' until a combination was found yielding the tristimulus values of the sample within the required accuracy.

The procedure which was followed obviated these difficulties. While at first it may seem to be unnecessarily cumbersome, in actual practice, it permitted the numerical integrations to be carried out in a systematic manner with satisfactory rapidity. In addition, the method was very convenient for the specification of homologous pairs. The basic principle of the procedure

TABLE I. Tristimulus specifications of constituent samples.

Test sample	Tristimulus values X	Y	Z
E_e	0.9804	1.0000	1.1812
$R1$	.1072	.0510	.0001
$R1v$	.1082	.0511	.0051
$Y1$	.4170	.3895	.0460
$Y1v$	.4180	.3896	.0511
$G1$	.3051	.4632	.0951
$G2$	.1777	.3312	.3324
$G2r$	.1799	.3320	.3324
$C1$	.1850	.2881	.5380
$C1r$	.1872	.2889	.5380
$B1$	.0727	.0275	.4077
$B1r$	.0750	.0284	.4077

was to define a preliminary set of twelve color samples as to spectral reflectance and a second set of eleven samples, by giving their effective equivalent densities. Each sample of this second set was then considered as an additive color combination of three members of the first set in order to fix its spectral reflectance curve. The members of the first and second sets we shall call constituent samples and composite samples, respectively. By way of physical analogy we may think of the composite sample as being produced on a Maxwell color disk by mixing the required amounts of three of the constituent colors.

Each of the constituent samples was specified by a set of selected wave-lengths* which represented the product of its spectral reflectance and the spectral energy distribution of Illuminant *C*. When a spectral sensitivity curve was summed at these wave-lengths the result was the value of the exposure as given by one of Eqs. (4). Evaluation of the tristimulus values of each constituent sample was carried out by summing the values of $\bar{x}$, $\bar{y}$, and $\bar{z}$ at the selected wave-lengths. The constituent samples are listed in Table I, together with their tristimulus specifications; their locations on the chromaticity diagram are shown in Fig. 7.

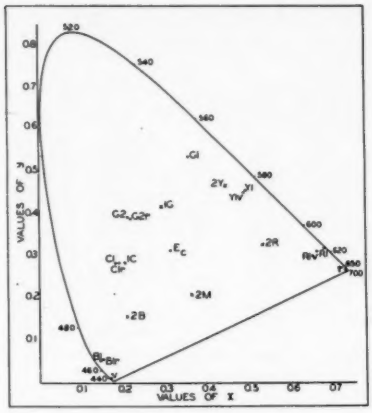

FIG. 7. Chromaticity diagram of color samples; constituent samples · · ·, composite samples x x x.

TABLE II. Specification of composite samples.

Composite sample	Effective equivalent densities j'	m'	c'	Tristimulus values X	Y	Z
$2R$ / $2Rv$	1.70	1.70	0.20	0.1425	0.0885	0.0381
$2Y$ / $2Yv$	2.20	0.20	0.20	0.4421	0.4637	0.0891
$1G$ / $1Gr$	1.00	0.10	1.00	0.2547	0.3687	0.2563
$1C$ / $1Cr$	0	0	1.50	0.2917	0.4098	0.7311
$2B$ / $2Br$	0.15	1.50	1.50	0.0586	0.0440	0.1781
$2M$	0.20	1.70	0.20	0.2948	0.1704	0.3467

* An error of nomenclature which has often appeared in the literature consists in referring to a table of wave-lengths as "selected ordinates." We propose the use of the phrase "selected wave-lengths" in this sense. The expression "selected ordinates" should be used to imply the values of a function at the selected wave-lengths.

The effective equivalent densities which characterize the composite samples were selected in such a way as to yield a suitable range of saturated colors. The corresponding tristimulus values were computed directly from the values of j', m', and c', using Eqs. (6). The effective equivalent densities and tristimulus specifications are listed in Table II while the corresponding points are located in the chromaticity diagram of Fig. 7. The samples consist of a main group of six which were designated $2R$, $2Y$, $1G$, $1C$, $2B$, and $2M$; supplementing these are five others, $2Rv$, $2Yv$, $1Gr$, $1Cr$, and $2Br$, each of which is identical in color with the corresponding member of the first group, but which has either some considerable far violet reflectance (denoted by "v") or some considerable far red reflectance (denoted by "r").

Referring again to Fig. 7 on which the chromaticity data for both sets of samples are plotted, we note that each composite sample falls within the triangle formed by three constituent samples, one of which may, in each case, be E_c. With the tristimulus values of all samples as given in Tables I and II it is a simple algebraic problem to determine the amounts of three constituent samples required to synthesize any composite sample which lies within their triangle. It now becomes clear how this method circumvents the difficulties of the other methods. Since the energy distributions of composite samples are obtained

TABLE III. Formulation of composite samples. $UR=k_1UR_1+k_2UR_2+k_3UR_3$.

Composite sample	Constituent samples			Coefficients		
	No. 1	No. 2	No. 3	k_1	k_2	k_3
2R	Y1	R1	E_e	0.0194	0.9676	0.0314
2Rv	Y1v	R1v	E_e	0.0353	0.9379	0.0268
2Y	G1	Y1	E_e	0.2744	0.8080	0.0219
2Yv	G1	Y1v	E_e	0.2802	0.8114	0.0178
1G	G2	G1	E_e	0.3738	0.3478	0.0838
1Gr	G2r	G1	E_e	0.3789	0.3465	0.0825
1C	B1	C1	E_e	0.0508	1.1532	0.0762
1Cr	B1r	C1r	E_e	0.0458	1.1686	0.0709
2B	B1	C1	E_e	0.3360	0.0001	0.0348
2Br	B1r	C1r	E_e	0.3352	0.0036	0.0335
2M	R1	B1	E_e	1.0655	0.4105	0.0424

by adding fractions of constituent samples, we may write the composite sample exposures as additive combinations of the exposures of corresponding constituent samples. The facility with which the graphical integrations could be carried out was an important factor in this choice of approach.

VII. COMPUTATIONAL PROCEDURE

The tristimulus values of the constituent samples were calculated using graphical plots of $\bar{x}$, $\bar{y}$, and $\bar{z}$ upon which the selected wave-length charts were superimposed. Using these values, the data for formulation of each composite sample from the three appropriate constituent samples were calculated by simultaneous solution of the following equations for k_1, k_2, and k_3:

$$\begin{aligned} X' &= k_1X_1+k_2X_2+k_3X_3, \\ Y' &= k_1Y_1+k_2Y_2+k_3Y_3, \\ Z' &= k_1Z_1+k_2Z_2+k_3Z_3, \end{aligned} \tag{7}$$

where X', Y', and Z' are the tristimulus values of the composite sample and X_1, Y_1, Z_1, X_2, Y_2, etc., are those of the constituent samples. The general formula for calculating UR, the spectral energy distribution from a composite sample of spectral irradiance U and spectral reflectance R, is then given by*

$$UR=k_1UR_1+k_2UR_2+k_3UR_3, \tag{8}$$

where R_1, R_2, and R_3 are the spectral reflectances of the three constituent samples of which the composite sample is composed. The data on the formulation of the composite samples are given in Table III which lists the values of the constants of Eq. (8) together with the constituent samples used.

In order to calculate the exposure to a given composite sample for a spectral sensitivity combination, each sensitivity component was integrated using the selected wave-length charts of the corresponding three constituent samples. Combining Eq. (8) with the first of Eqs. (4) we have, for the exposure E_b, of a color film layer of sensitivity S_b, to the composite sample of spectral reflectance R,

$$E_b=\int_0^\infty URS_bd\lambda=k_1\int_0^\infty UR_1S_bd\lambda + k_2\int_0^\infty UR_2S_bd\lambda+k_3\int_0^\infty UR_3S_bd\lambda. \tag{9}$$

Considering S_b to be a spectral sensitivity component, we were able to integrate each term of the right member of Eq. (9) using the selected wave-length charts which specify UR_1, UR_2, and UR_3. Where a combination of two or more spectral sensitivity components were under consideration, the separate exposures were weighted and added according to the composition of the spectral sensitivity curve. Finally, using Eqs. (5), the equivalent densities of each reproduction were calculated with results as listed in Tables IV–VI. These tables give the effective equivalent densities for each composite sample, together with the corresponding equivalent densities of the reproductions.

As an illustration of the application of the method we may consider the calculation of the equivalent density value constituting the reproduction of the green composite sample 1G by the spectral sensitivity $R5^403^402^4$. This sensitivity which is shown in Fig. 8 consists of a combination of three sensitivity components each of forty millimicrons width located with maxima at 620, 640, and 660 mμ and of relative heights 5, 3, and 2. From the equivalent densities $j=1.00$, $m=0.10$, and $c=1.00$ corresponding to perfect reproduction of the composite sample, the tristimulus values were calculated by use of Eq. (6). Each of the two constituent samples $G1$ and $G2$ were formulated by choosing a smoothly spaced set of selected

* The spectral irradiance U is retained in Eq. (8) since the selected wave-length charts, which were used in evaluating Eqs. (3) and (4), specified in each case the product of the spectral energy and the spectral reflectance of the constituent sample.

TABLE IV. Reproduction of composite samples by a series of blue sensitivity combinations.

Sensitivities	Samples										
	2R	*2Rv*	*2Y*	*2Yv*	*1G*	*1Gr*	*1C*	*1Cr*	*2B*	*2Br*	*2M*
	$j'=1.70$	1.70	2.20	2.20	1.00	1.00	0.00	0.00	0.15	0.15	0.20
					Reproductions						
$B1^4 0.0$	$j=1.80$	1.69	1.99	1.87	1.03	1.03	0.32	0.32	0.84	0.84	0.73
$B1^4.0$	$j=1.80$	1.76	1.91	1.88	0.94	0.94	0.28	0.28	0.84	0.84	0.74
$B0.1^4$	$j=1.79$	1.80	1.57	1.58	0.86	0.86	0.25	0.25	0.86	0.86	0.76
$B0.01^4$	$j=1.79$	1.81	1.22	1.24	0.76	0.76	0.23	0.23	0.90	0.90	0.79
$B0.001^4$	$j=1.78$	1.82	1.04	1.04	0.67	0.67	0.20	0.20	0.95	0.95	0.84
$B0.0001^4$	$j=1.76$	1.81	0.95	0.95	0.61	0.60	0.19	0.19	1.02	1.02	0.91
$B1^2.0$	$j=1.80$	1.79	1.99	1.98	1.08	1.08	0.30	0.30	0.86	0.86	0.75
$B0.1^2$	$j=1.80$	1.84	1.99	2.03	0.95	0.95	0.26	0.26	0.86	0.86	0.76
$B0.01^2$	$j=1.80$	1.88	1.55	1.58	0.72	0.72	0.23	0.23	0.86	0.86	0.76
$B5^4 0.5^2$	$j=1.80$	1.76	1.99	1.94	0.98	0.98	0.29	0.29	0.85	0.85	0.75
$B3^4 0.7^2$	$j=1.80$	1.79	1.99	1.98	0.97	0.97	0.28	0.28	0.85	0.85	0.76

TABLE V. Reproduction of composite samples by a series of green sensitivity combinations.*

Sensitivities	Samples										
	2R	*2Rv*	*2Y*	*2Yv*	*1G*	*1Gr*	*1C*	*1Cr*	*2B*	*2Br*	*2M*
	$m'=1.70$	1.70	0.20	0.20	0.10	0.10	0.00	0.00	1.50	1.50	1.70
					Reproductions						
$G1^2 0.0$	$m=1.70$	1.69	0.41	0.41	0.37	0.37	0.24	0.24	1.70	1.70	1.61
$G1^2.0$	$m=1.70$	1.69	0.41	0.41	0.38	0.38	0.30	0.30	1.74	1.74	1.63
$G0.1^2$	$m=1.69$	1.68	0.38	0.38	0.43	0.43	0.38	0.38	1.75	1.75	1.64
$G0.01^2$	$m=1.68$	1.64	0.32	0.32	0.46	0.46	0.48	0.48	1.75	1.75	1.64
$G1^4 0.0$	$m=1.72$	1.72	0.48	0.48	0.40	0.40	0.25	0.25	1.61	1.61	1.50
$G1^4.0$	$m=1.70$	1.69	0.41	0.41	0.41	0.41	0.30	0.30	1.67	1.67	1.57
$G0.1^4$	$m=1.69$	1.68	0.38	0.38	0.43	0.43	0.38	0.38	1.72	1.72	1.61
$G0.01^4$	$m=1.68$	1.67	0.36	0.36	0.48	0.48	0.42	0.42	1.74	1.74	1.63
$G1^6.0$	$m=1.70$	1.70	0.44	0.44	0.43	0.43	0.31	0.31	1.57	1.57	1.46
$G0.1^6$	$m=1.69$	1.68	0.40	0.40	0.46	0.46	0.37	0.37	1.67	1.67	1.57
$G0.1^1$	$m=1.70$	1.68	0.38	0.38	0.44	0.44	0.38	0.38	1.75	1.75	1.64

* In the case of the green series, it was found necessary to test only individual components.

TABLE VI. Reproduction of composite samples by a series of red sensitivity combinations.

Sensitivities	Samples										
	2R	*2Rv*	*2Y*	*2Yv*	*1G*	*1Gr*	*1C*	*1Cr*	*2B*	*2Br*	*2M*
	$c'=0.20$	0.20	0.20	0.20	1.00	1.00	1.50	1.50	1.50	1.50	0.20
					Reproductions						
$R1^4 000.0$	$c=0.79$	0.79	0.29	0.29	0.80	0.80	1.25	1.27	1.75	1.77	0.76
$R1^4 00.0$	$c=0.72$	0.72	0.29	0.29	0.88	0.88	1.34	1.37	1.75	1.76	0.67
$R1^4 0.0$	$c=0.61$	0.61	0.30	0.30	1.00	1.00	1.34	1.36	1.75	1.75	0.56
$R1^4.0$	$c=0.48$	0.48	0.30	0.30	1.04	1.03	1.34	1.27	1.75	1.70	0.43
$R0.1^4$	$c=0.41$	0.42	0.31	0.31	1.09	1.06	1.34	1.19	1.75	1.64	0.36
$R0.01^4$	$c=0.35$	0.36	0.34	0.34	1.21	1.15	1.34	1.13	1.75	1.61	0.30
$R0.001^4$	$c=0.31$	0.32	0.36	0.36	1.27	1.16	1.34	1.02	1.75	1.54	0.26
$R5^4 003^4.02^4$	$c=0.58$	0.58	0.30	0.30	0.94	0.92	1.30	1.24	1.75	1.72	0.53
$R6^4 03^4.01^4$	$c=0.59$	0.60	0.30	0.31	0.95	0.95	1.34	1.31	1.75	1.72	0.55
$R5^4 03^4.02^4$	$c=0.55$	0.55	0.30	0.30	0.98	0.97	1.34	1.28	1.75	1.72	0.50
$R5^4 02^4.03^4$	$c=0.53$	0.54	0.30	0.30	0.99	0.98	1.34	1.27	1.75	1.70	0.49
$R4^4 02^4.2^4 02^4$	$c=0.50$	0.50	0.31	0.31	1.01	1.00	1.34	1.22	1.75	1.68	0.46
$R4^4 01^4.2^4 03^4$	$c=0.49$	0.49	0.31	0.31	1.03	1.00	1.34	1.20	1.75	1.66	0.43
$R3^4 02^4.2^4 03^4$	$c=0.47$	0.47	0.31	0.31	1.04	1.02	1.34	1.19	1.75	1.66	0.42

wave-lengths which yielded suitable trichromatic coefficients. The tristimulus values of the constituent samples were determined by summation of the selected ordinates of the I.C.I. tristimulus value functions and from them, together with the tristimulus values of E_e, the constants k_1, k_2, and

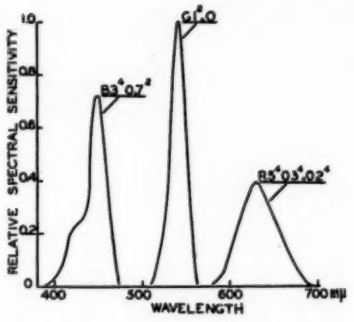

FIG. 8. Optimum spectral sensitivity curves for a reversible subtractive color process employing primaries of Fig. 3.

k_3 were calculated by use of Eq. (7). The red spectral sensitivity component of 40 mμ width was integrated with its maximum position first at 620 mμ using successively the three selected wave-length charts for $G1$, $G2$, and E_c; then the integrations were repeated using the same curve and charts, but with sensitivity maxima at 640 mμ and 660 mμ. These nine integrated values, according to Eq. (4) constituted the exposures of each of the sensitivity components to each of the constituent samples. The exposure values were multiplied by the appropriate values of k_1, k_2, and k_3 and added to yield the three exposures of the composite sample $1G$ to these same sensitivity components. Finally the exposure values were respectively multiplied by 0.5, 0.3, and 0.2 and added yielding the exposure, the logarithm of which multiplied by -1.2 according to Eq. (5), gave the equivalent density $c=0.98$ as listed (underlined) in Table VI.

VIII. ANALYSIS OF DATA FOR OPTIMUM SENSITIVITIES

The final evaluation of the spectral sensitivity functions which were analyzed depends upon the comparison between the reproduction equivalent densities j, m, and c and the effective equivalent densities j', m', and c' which characterize the color samples. The data as presented in Tables IV, V, and VI compare the results for a total of thirty-six spectral sensitivity functions.

Four important criteria of optimum color rendition may be validly used in the selection of the optimum spectral sensitivities. The first two of these are related to the proper rendition of homologous pairs and the need for maximum lightness and saturation. The third criterion is concerned with hue rendition and therefore provides a basis for judging the performance of triple sets of sensitivities. The final decision among near-optimum sensitivities is made on the practical consideration of their probable relative speeds.

Applying these four tests in succession, the first basis on which to judge the various sensitivities is to compare their responses to pairs of color samples which are colorimetrically identical but spectrophotometrically different. For example, of the series of blue sensitivities $B1^40.0$, $B1^4.0 \ldots\ldots B0.0001^4$, with maxima at 430, 440 . . . . . 480 mμ the one at 450 mμ reproduces the two red samples most nearly the same, as it does also the pair of yellow samples. It is significant to note that this maximum is very near the peak of the I.C.I. tristimulus value function $\bar{z}$ at 445 mμ.

In addition, in order to achieve sufficient lightness and saturation, we require that, with a sample whose effective equivalent densities are large, the equivalent densities of the reproduction should be a maximum, and that where the effective equivalent densities are small, the reproduction equivalent densities should be a minimum. In this respect, $B3^40.7^2$ is to be preferred over $B0.1^4$ because of the higher density obtained in the green reproduction ($j=0.97$ compared with $j=0.86$). The third criterion, which implies consideration of blue, green, and red sensitivities in combination and is related to the rendition of hue, is that, for samples having two effective equivalent densities which are equal, the corresponding two equivalent densities of the reproduction would be approximately equal. This condition depends upon the fact that each of the dyes under consideration remains reasonably constant in dominant wave-length as we change its density over the photographically usable range. As an example, consider the renditions of sample $1G$ by $B1^2.0$ and by $B3^40.7^2$. With these two sensitivities, the values of j which are 1.08 and 0.97, respectively, are to be compared with the value $c=0.98$ for the optimum red sensitivity. $B3^40.7^2$ is to be preferred over $B1^2.0$ since the latter sensitivity yields a relatively high equivalent density value; the resulting reproduction

would, therefore, be significantly more yellow than the sample.

A fourth consideration is, that of two sensitivities which give nearly the same results, the one with the wider curve is to be preferred. The principle reason lies in the fact that a layer having a broader sensitivity may usually be expected to be photographically faster. In addition it will be less liable to yield hue distortion as a result of small variations in sample reflectance near the wave-length of maximum sensitivity. As an application of this criterion, we may again refer to the two blue sensitivities $B1^{2}.0$ and $B3^{4}0.7^{2}$, the former of which is much more narrow and therefore less acceptable.

Using these criteria as a basis, the data of Tables IV–VI were analyzed, yielding optimum spectral sensitivity curves as shown in Fig. 8. It was found that small variations of these curves produced relatively small changes in the corresponding reproductions, while larger variations resulted, in each case, in over-all increase of color distortion. The three curves thus represent, as functions of wave-length, the optimum sensitivities of the three layers of a reversible subtractive color process, assuming that the process employs dyes whose spectral characteristics are shown in Fig. 3.

IX. EVALUATION OF EXPERIMENTAL SENSITIZERS

The procedure described above permitted us to derive the optimum sensitivities which serve as goals in the development of a given color process. In practice, however, we are also often interested in determining which of two available sensitivities is to be preferred with respect to its color rendition characteristics. The manner in which such an analysis is carried out may be illustrated by calculating the renditions of two red sensitivities assuming each to be used with type blue and green sensitivities. The curves of the type* sensitivities are shown in solid lines in Fig. 9 while that of the comparison red sensitivity is indicated in dotted lines.

The comparison was carried out by computing, for each sensitivity curve, the corresponding equivalent density values. In this case, of course, we had the curve of each sensitivity so that the use of component sensitivities was unnecessary. The integrations of the type $\int_0^\infty URSd\lambda$ were carried out for each sensitivity curve, used together with each constituent color sample. The resulting exposures were weighted according to the constants of Table III to give the exposures for the eleven composite samples. Substituting in Eqs. (5) which were

$$\begin{aligned} j &= -\gamma \log E_b, \\ m &= -\gamma \log E_g, \\ c &= -\gamma \log E_r, \end{aligned} \tag{5}$$

we obtained the equivalent densities as listed in Table VII. For comparison purposes, we have also tabulated the corresponding values for optimum reproductions.

We may first discuss the errors in the type reproduction as compared with those for optimum sensitivities in order to see the directions in which greatest improvements might be expected; following that we will examine the data on the comparison red sensitivity to see how it compares with type.

In one outstanding respect, the type material gives better rendition than optimum—that is in the reproduction of the reds. In this case the cyan density is much lower than the corresponding value for the optimum sensitivities due to the fact that the red sensitivity has its maximum at longer wave-lengths. This improvement is achieved, however, at the cost of severe increases in distortion of the renditions of other colors.

In the other cases, the type sensitivities gave considerably poorer than optimum reproductions.

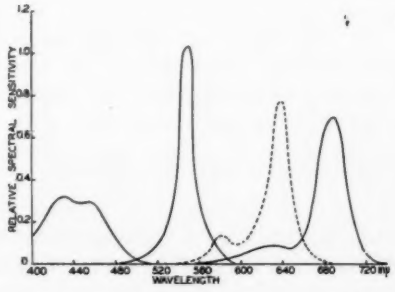

FIG. 9. Spectral sensitivity curves; type blue, green and red sensitivities ————————, comparison red sensitivity – – – – – – – – – – –.

* The type sensitivities were those which characterized an experimental color film made prior to the war. The comparison red sensitivity was one of a series which was developed in the course of research directed toward improved color rendition.

TABLE VII. Reproductions of composite samples by type and comparison red spectral sensitivities.

Sample number		Sample equiv. dens.	Reproduction equiv. dens. Type	Comparison	Optimum
	j	1.70	1.79	1.79	1.80
2*R*	*m*	1.70	1.70	1.70	1.70
	c	0.20	0.29	0.61	0.55
	j	1.70	1.73	1.73	1.79
2*Rv*	*m*	1.70	1.68	1.68	1.69
	c	0.20	0.30	0.62	0.55
	j	2.20	1.51	1.51	1.99
2*Y*	*m*	0.20	0.40	0.40	0.41
	c	0.20	0.34	0.30	0.30
	j	2.20	1.49	1.49	1.98
2*Yv*	*m*	0.20	0.40	0.40	0.41
	c	0.20	0.34	0.30	0.30
	j	1.00	0.91	0.91	0.97
1*G*	*m*	0.10	0.42	0.42	0.38
	c	1.00	1.18	0.88	0.98
	j	1.00	0.91	0.91	0.97
1*Gr*	*m*	0.10	0.42	0.42	0.38
	c	1.00	1.03	0.88	0.97
	j	0	0.29	0.29	0.28
1*C*	*m*	0	0.37	0.37	0.30
	c	1.50	1.28	1.14	1.34
	j	0	0.29	0.29	0.28
1*Cr*	*m*	0	0.37	0.37	0.30
	c	1.50	0.86	1.13	1.28
	j	0.15	0.89	0.89	0.85
2*B*	*m*	1.50	1.72	1.72	1.74
	c	1.50	1.75	1.75	1.75
	j	0.15	0.89	0.89	0.85
2*Br*	*m*	1.50	1.72	1.72	1.74
	c	1.50	1.43	1.74	1.72
	j	0.20	0.77	0.77	0.76
2*M*	*m*	1.70	1.62	1.62	1.63
	c	0.20	0.24	0.56	0.50

The yellows were not nearly as saturated, as indicated in the case of sample 2*Y* by a yellow density of 1.51 as compared with an optimum value of 1.99. The greens suffered, in that the values of yellow and cyan densities for the reproduction of 1*G* were 0.91 and 1.18, respectively. Thus, the reproduction would be more cyan than the sample. In addition, the fact that the two values of cyan density for samples 1*G* and 1*Gr* are considerably different shows that two greens of identical color would be reproduced differently if one of them were infra-red reflecting. This latter result may be directly attributed to the spectral location of the red sensitivity. The cyans and blues suffered for the same reason. The magenta reproduction was found to be much too red as indicated by yellow and cyan density values of 0.77 and 0.24, respectively.

On the basis of the above analysis, we may conclude that the greatest single cause of poor reproduction is the position of the red sensitivity. The comparison sensitivity with a maximum at 630 mμ may, therefore, be expected to show improvement.

Comparison of the data proves that this is indeed the case although there is at the same time strong evidence that the secondary maximum at 580 mμ introduces other errors. These show up principally in the renditions of the reds, greens, and cyans. In each of these cases the homologous pairs were reproduced nearly identically as regards values of *c*. Each of the reds, however, had too high a cyan density, while the other two pairs of colors showed cyan densities which were lower than optimum because of the secondary red maximum. The reproductions of the two blues as well as those of the magentas were nearly optimum.

Summarizing the above remarks, we may say that the comparison red sensitivity is much better than type but would be further improved by eliminating the secondary maximum at 580 mμ. It would also be desirable to extend the sensitivity to longer wave-lengths in order to conform more closely with the optimum curve. This would not only tend to balance out small reflectance variations in this wave-length band but might also be expected to result in greater emulsion speed for the red sensitive layer.

In addition to forming a basis for a critical evaluation of type and comparison red sensitivities, the data of Table VII permit us to compare type blue and green sensitivities with optimum. The outstanding error which may be attributed to them is found in the case of the low saturation of the reproductions of the yellow samples. This result is because of the fact that the blue sensitivity is too high in the vicinity of 500 mμ. In addition some indication of too high an ultraviolet sensitivity is shown by the comparative values of *j* for reproductions of 2*R* and 2*Rv*.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 37, NUMBER 2 MARCH, 1947

A Double-Beam, Percent Transmission Recording Infra-Red Spectrophotometer

N. WRIGHT AND L. W. HERSCHER
The Dow Chemical Company, Midland, Michigan
(Received January 8, 1947)

A spectrophotometer for the infra-red is described which consists of an external photometering system adapted to an ordinary rocksalt prism spectrometer equipped with a fast responding bolometer. Two radiation beams from the same source are made to fall alternately (at 6 cycles per sec.) upon the slit of the spectrometer while a variable diaphragm automatically balances one of the beams against the other. Electronic amplification and pen-type recording are employed to register the spectra, which appear directly as percent of radiation transmitted. Differential comparison of two samples, one placed in each beam, is an important capability of the instrument. When impure and pure samples are thus compared the spectrophotometer automatically records the spectrum of the impurity only. Similarly in the case of solution and solvent the spectrum of solute only is recorded. Recordings showing the performance of the apparatus are reproduced.

INTRODUCTION

INFRA-RED spectroscopists have long recognized the desirability of a spectrophotometer capable of direct recording of percent of radiation transmitted *versus* wave-length. That the instrument should also have the capability of comparing differentially the spectra of two samples has been regarded as of equal importance, particularly by those applying the spectra for chemical analysis.

These advantages were realized for the visible range many years ago in the excellent recording spectrophotometer of A. C. Hardy,[1,2] and the soundness of the principle of null-balancing of two radiation beams has been well demonstrated by the success of that instrument.

The first successful use of beam balancing in a spectrophotometer for infra-red was made by J. D. Hardy and A. I. Ryer[3] in 1939. The balancing operation in this case was performed manually and, hence, speed and accuracy could not be achieved simultaneously. In 1942, Lehrer[4] described an infra-red spectrophotometer employing automatic balancing of radiation beams similar in some respects to the instrument presently to be described. As in the first case[3] the two beams took slightly different paths through the dispersing optics, entering the slit one above the other. Other disadvantages in the optics and electro-mechanics of photometering were evidenced in some instability and apparent non-linearity of the recordings. Another, quite different type of spectrophotometer has been described by E. B. Baker and C. D. Robb,[5] but this did not employ direct balancing of the radiation beams.

In starting the development of the spectrophotometer of this report it appeared at once that null-balancing of the two radiation beams would eliminate a large number of the difficulties encountered in infra-red spectroscopy. The absorption bands caused by last traces of water vapor and CO_2 in the optical path would be cancelled, as would all the minor absorptions resulting from windows, the prism, etc., traversed by both beams. Emission and absorption characteristics of the radiation source and bolometer receiver, respectively, would have no effect, as would be true also of the characteristics (non-linearity, variation in gain, etc.) of the bolometer amplifier. It was also recognized that periodic alternation or "flickering" of the two beams would eliminate "zero drift" and at the same time allow use of the more stable a.c. type electronic amplifier.

SPECTROMETER PROPER AND BOLOMETER

The spectrometer proper is a rocksalt prism instrument of the conventional type, a diagram of which is shown in Fig. 1. It employs a Littrow arrangement, the light passing twice through a

[1] A. C. Hardy, J. Opt. Soc. Am. and Rev. Sci. Inst. **18**, 96 (1929).
[2] A. C. Hardy, J. Opt. Soc. Am. **25**, 305 (1935).
[3] J. D. Hardy and A. I. Ryer, Phys. Rev. **55**, 1112 (1939).
[4] E. Lehrer, Zeits. f. tech. Physik **23**, 169 (1942).
[5] E. B. Baker and C. D. Robb, Rev. Sci. Inst. **14**, 362 (1943).

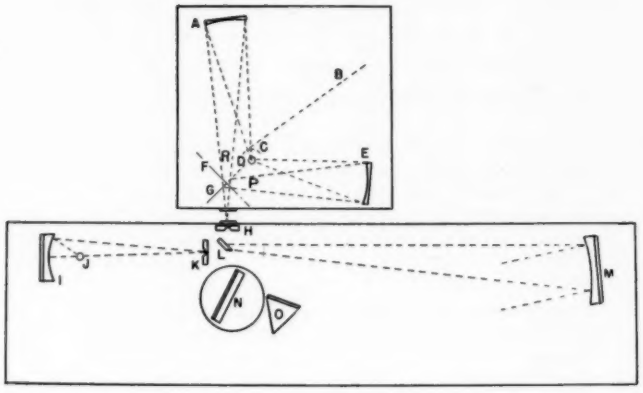

FIG. 1. Optical diagram of photometer (upper box) and spectrometer. *D*—source, *A* and *E*—focusing mirrors, *C*—comb diaphragm, driven by worm and wheel, *B*—worm shaft to balancing motor of recorder, *G*—rotating sector mirror, *F*—axis of sector mirror, *P* and *R*—locations of sample and comparison sample, respectively.

60° prism of 10-×8-cm faces. The collimator has a focal length of 92 cm, and the ellipsoidal mirror *I*, which focuses the radiation on the detector *J*, has foci at 37.5 and 7.5 cm, respectively.

The only unusual feature within this part of the spectrometer is the radiation detector, which is a bolometer capable of responding to radiation modulated at 6 c.p.s. It consists of a nickel strip, 10 mm×0.5 mm×0.1 micron (approx.), produced by the method of Brockman,[6] upon which a metallic black of Au or Ag is deposited. Its electrical resistance is in the neighborhood of 20 ohms. The bolometer is operated at a pressure of about 10^{-1} mm Hg, and is activated by 0.4 to 0.6 volt from a storage battery, as shown in the circuit diagram, Fig. 2. The output is applied to the primary of the transformer T_1 using the condenser circuit shown. (A bridge circuit may, however, be substituted with little loss.) The sensitivity is .1 microvolt (6 cycle r.m.s. at the primary of T_1) per micro watt of radiation.

PHOTOMETER OPTICS

In the external optical system two beams of radiation from a Globar source are made to fall alternately on the entrance slit of the spectrometer by means of a rotating sector-mirror in the manner illustrated in Fig. 1. The sector

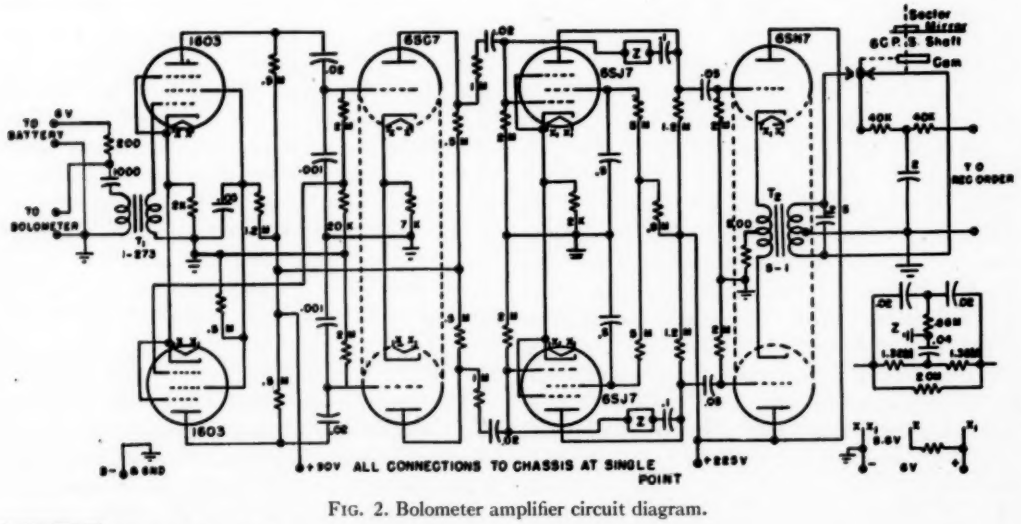

FIG. 2. Bolometer amplifier circuit diagram.

[6] F. G. Brockman, J. Opt. Soc. Am. **36**, 32 (1946).

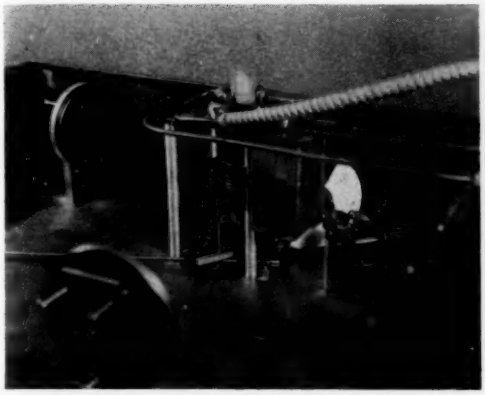

FIG. 3. View of photometer showing semi-circular sector-mirror and comb-diaphragm. The width of the comb is 2.5 cm and the tooth length is 2.5 cm.

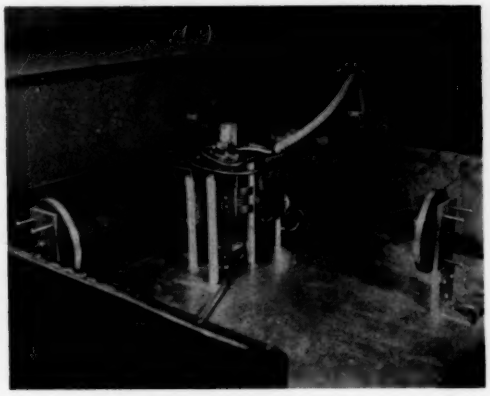

FIG. 4. Different view of photometer showing the two wells which hold the sample cells.

mirror G (see also Fig. 3) is semicircular in shape and rotates (on shaft F, Fig. 1) at 6 c.p.s. One beam passes through the sample (at P) while the other passes first through a variable attenuating diaphragm (C), then through a dummy cell or standard sample (at R). When the rotating sector-mirror is above the shaft the sample beam is reflected into the spectrometer and the standard beam is blocked off; when the mirror has rotated below the shaft the standard beam enters the spectrometer and the sample beam passes by and is lost.

The variable attenuating diaphragm, which automatically balances the intensity of the standard beam against that of the sample beam, has the form of a comb, with triangular teeth (Figs. 3 and 4). It is mounted vertically at a position near the source and is driven horizontally across the beam by means of a worm gear arrangement connected to the balancing motor of a pen recorder.* Thus there is a linear mechanical coupling between pen and diaphragm motions. The comb itself, instead of the globar, is imaged on the entrance slit and the spectrometer aperture is filled out at all comb positions.

The mechanism of automatic balancing of the two beams is best described by considering that, for the moment, the beams are unequal. Thus there will be a 6-c.p.s. rise and fall of the bolometer temperature and a consequent output potential difference pulsating at this frequency. The phase is determined by whichever of the two beams is the more intense. This 6-cycle component of the bolometer output is amplified by a special transformer** and the four-stage tuned feed-back amplifier shown in Fig. 2. Full-wave rectification of the a.c. output is effected by a cam switch mounted on the sector-mirror shaft. The resulting d.c. output, of which the polarity is now determined by whichever of the beams is the more intense, passes to the recorder where, after again being amplified, it actuates the balancing motor (Fig. 5). In turn the balancing

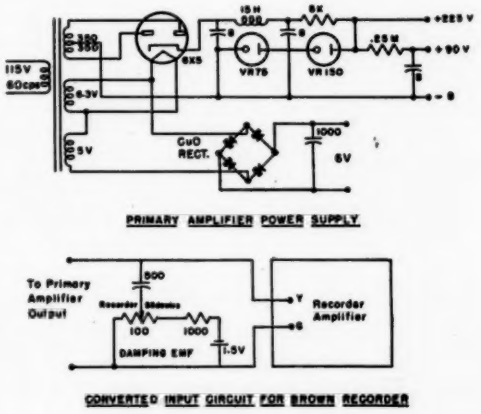

FIG. 5.

* Brown Instrument Company. Model No. 153×12P ×30.

** United Transformer Corporation, No. 86300.

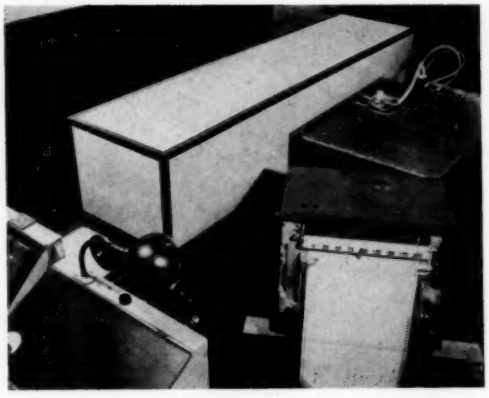

FIG. 6. Photograph of the spectrophotometer.

motor moves the comb-diaphragm in the proper direction to balance the two radiation beams, simultaneously moving the recording pen to the new position corresponding with that of the diaphragm.

The relationship between movement of the comb-diaphragm (hence recording pen position) and intensity of the balancing beam is linear to within 1 percent of full scale throughout the range from 0 to 100 percent transmitted. Some non-linearity near the 0 percent end was corrected by covering the extreme tips of the angles between the teeth with a straight edge. Further correction was obtained subsequent to making the photographs (Figs. 3 and 4) by increasing the number of teeth from four to six.

The spectrum is scanned in the usual manner by a uniform, motor driven, angular displacement of the Littrow mirror. A pair of Selsyn motors synchronizes with this the motion of the recorder paper. Also coupled to the wave-length drive is a linear gear mechanism for continuously varying the slit widths, so that the spectral energy arriving at the bolometer (with sample cell empty) remains roughly constant throughout the spectrum. Thus the instrument automatically records the percent of radiation transmitted by the sample *versus* wave-length (on a non-linear wave-length scale). The time of recording the range 2 to 15μ is 13 minutes. An over-all view of the apparatus is shown in Fig. 6.

PERFORMANCE

The performance of the apparatus is illustrated by the photographs of spectra shown in Figs. 7–10. In Fig. 7, the blank run (100 percent transmitted), made with no cell in either beam, exhibits the accurate balance obtained throughout the spectrum. One will notice that there is no trace of the 6μ band of water vapor nor of the CO_2 bands at 4.26μ and 13.9μ. The transmission spectrum of glass reveals some spectral impurity at wave-lengths greater than 6μ. This is a fault of the spectrometer itself which the double beam optics eliminates only in the case of differential

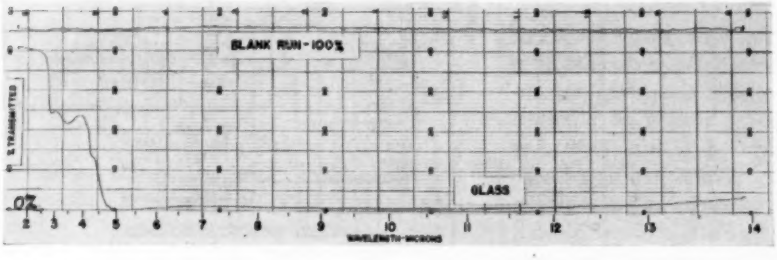

FIG. 7. Upper trace—blank run with no cell in either beam. Lower—transmission of glass, 1 mm thick. (Disregard numbers printed on the chart paper.)

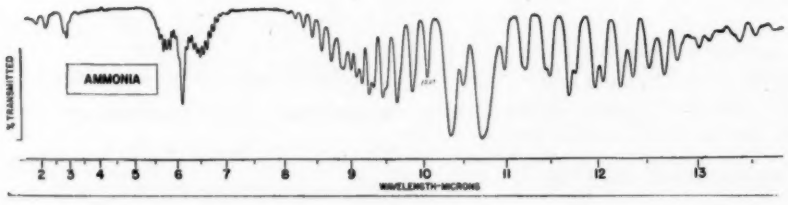

FIG. 8. Spectrum of ammonia vapor, repeated on same chart, 5-cm cell, pressure—225 mm Hg.

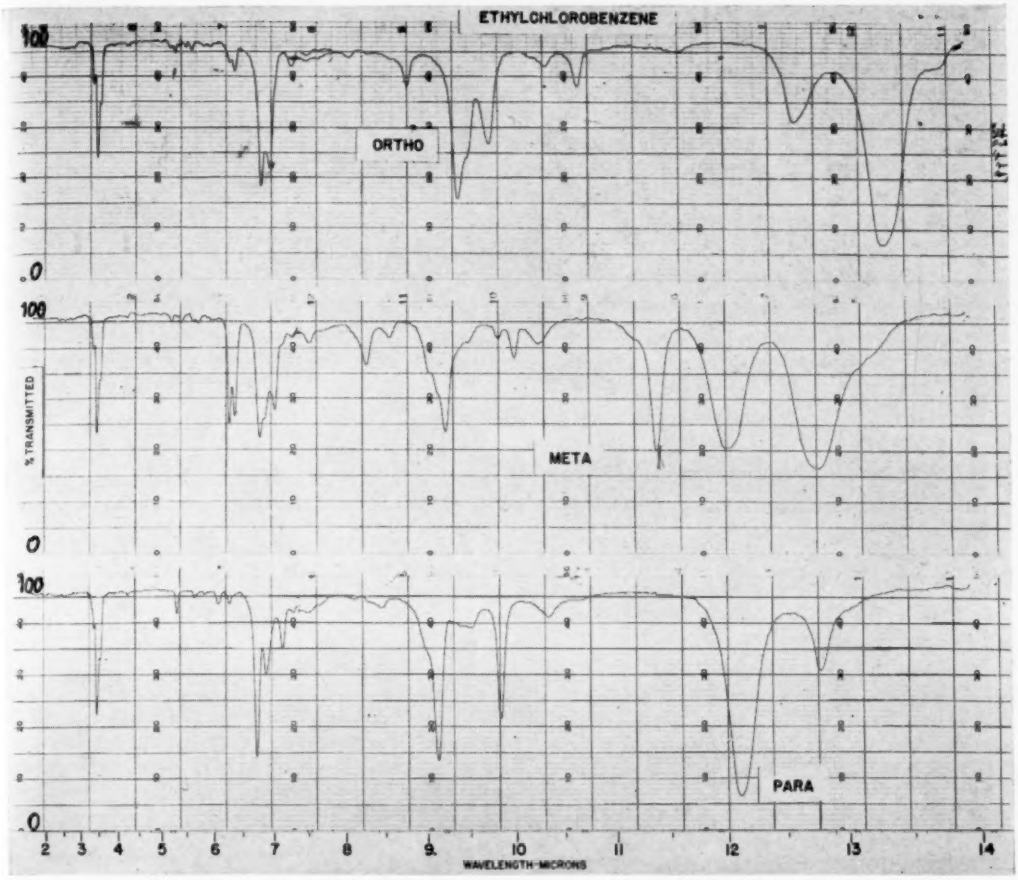

FIG. 9. Spectra of ortho-, meta-, and para-ethylchlorobenzene: 10 percent solution in CCl_4, 2 to 7.5μ, 10 percent solution in CS_2, 7.5 to 14μ. Thickness 0.1 mm. Absorption bands of solvent have been removed by placing pure solvent in comparison beam.

comparison of impure with pure samples. There is, of course, complete freedom from zero drift.

Figure 8, showing the spectrum of ammonia vapor, demonstrates the precision of the apparatus; after making one recording the instrument and chart were reset at the starting point and a second recording was superimposed upon the first. Irregularity evidenced in a few places by a doubling of the trace averages much less than 1 percent of full scale.

Some typical recordings of spectra of organic compounds are shown in Fig. 9. The samples were measured as 10 percent solutions in CS_2 and CCl_4, and in each case a second cell containing the corresponding pure solvent was placed in the comparison beam. Thus the absorption bands caused by the solvents were automatically removed from the spectra.

Use of the instrument as a differential spectrophotometer is best demonstrated in Fig. 10 showing how the spectrum of catechol is automatically obtained by placing an acetone solution of catechol in the sample beam, and an equivalent thickness of acetone in the comparison beam. It is to be noticed that the catechol spectrum is completely free of acetone bands. In regions where acetone absorbs totally the instrument is, of course, rendered inactive.

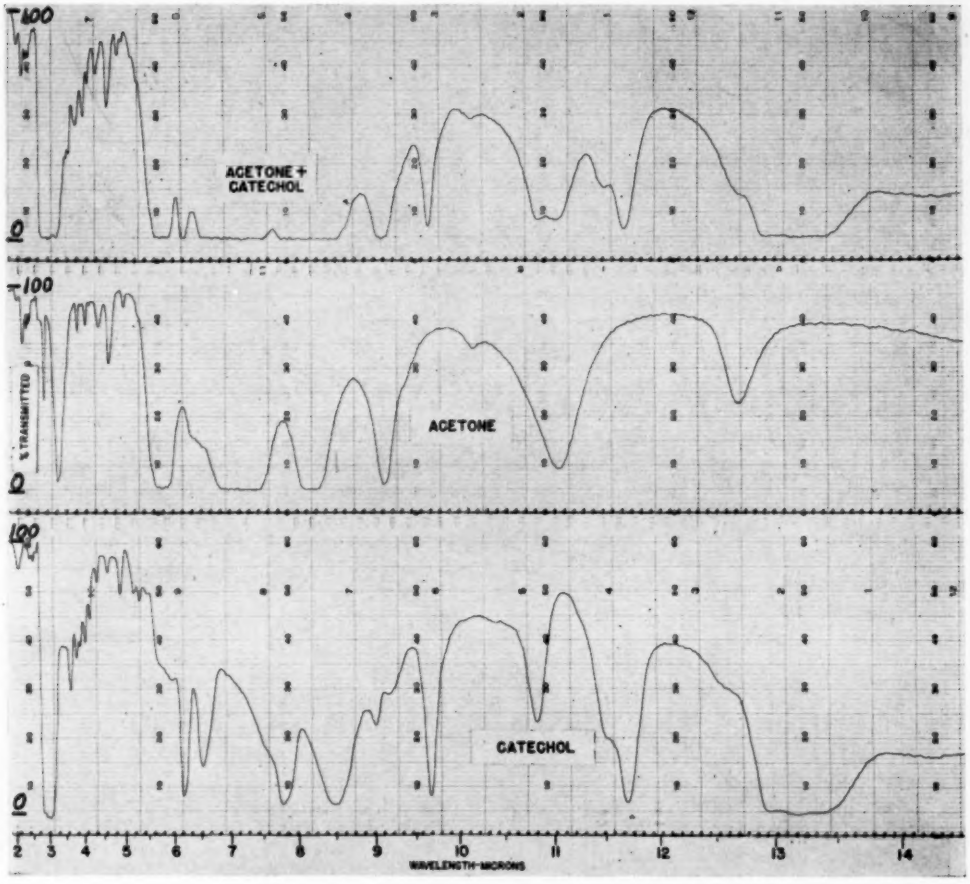

FIG. 10. Upper spectrum—solution of catechol in acetone, 19.5 g per 100-cc solution, cell thickness 0.117 mm. Center spectrum—pure acetone, 0.117-mm cell. Lower spectrum—sample beam contains catechol in acetone, 19.5 g per 100-cc solution, 0.117 mm; comparison beam contains pure acetone in 0.100-mm cell.

CONCLUSION

The spectrophotometer has come up to expectations regarding the important objectives. Some betterment of stability and signal-to-noise ratio is needed, and it is expected that this will be achieved through installation of a more sentive radiation detector. Most noteworthy of the many advantages of the instrument is that of differential comparison of two samples. The capability of recording the spectrum of a compound in solution while automatically cancelling out the spectrum of the solvent, broadens greatly the usefulness of the infra-red method. Analysis can thus be extended to the large class of compounds insoluble in the favored infra-red solvents, CS_2 and CCl_4, but soluble in acetone, alcohol, ether, etc. In particular this should facilitate the use of infra-red spectroscopy in the fields of biological and medical research.